BİLGİ TEKNOLOJİLERİ VE İLETİŞİM KURUMU

MAKİNE ÖĞRENMESİ KAPSAMINDA YAPAY ZEKA TEKNOLOJİLERİNİN SİBER GÜVENLİĞİN SAĞLANMASINDA KULLANIMI, ÜLKE UYGULAMALARI VE ÜLKEMİZ İÇİN ÖNERİLER

Doğukan Ömer GÜR

Bilişim Uzmanlık Tezi

Ağustos 2024

Ankara

BİLGİ TEKNOLOJİLERİ VE İLETİŞİM KURUMU

MAKİNE ÖĞRENMESİ KAPSAMINDA YAPAY ZEKA TEKNOLOJİLERİNİN SİBER GÜVENLİĞİN SAĞLANMASINDA KULLANIMI, ÜLKE UYGULAMALARI VE ÜLKEMİZ İÇİN ÖNERİLER

Doğukan Ömer GÜR

Bilişim Uzmanlık Tezi

Ağustos 2024

Ankara

Doğukan Ömer Gür tarafından hazırlanan Makine Öğrenmesi Kapsamında Yapay Zeka Teknolojilerinin Siber Güvenliğin Sağlanmasında Kullanımı, Ülke Uygulamaları ve Ülkemiz İçin Öneriler adlı bu tezin Bilişim Uzmanlığı tezi olarak uygun olduğunu onaylarım.

Bilişim Uzmanı Hasan Kağan AKICI
Tez Danışmanı

Bu çalışma, tez savunma komisyonumuz tarafından Bilişim Uzmanlığı tezi olarak kabul edilmiştir.

Başkan : Kurul Üyesi, Nurettin ŞAR

Üye : Daire Başkanı, Mahmut Esat YILDIRIM

Üye : Bilişim Uzmanı, Hasan Kağan AKICI

Üye : Bilişim Uzmanı, Ömer TUNÇ

Üye : Bilişim Uzmanı, Ömer Faruk YALLI

Bu tez, Bilgi Teknolojileri ve İletişim Kurumu tez yazım kurallarına uygundur.

ÖZET

BİLGİ TEKNOLOJİLERİ VE İLETİŞİM KURUMU	
Tezin Adı	Makine Öğrenmesi Kapsamında Yapay Zeka Teknolojilerinin Siber Güvenliğin Sağlanmasında Kullanımı, Ülke Uygulamaları ve Ülkemiz İçin Öneriler
Türü	Bilişim Uzmanlık Tezi
Yazar	Doğukan Ömer Gür
Teslim Tarihi	29.06.2024
Anahtar Kelimeler	Siber Güvenlik, Yapay Zeka, Makine Öğrenmesi
Tez danışmanı	Hasan Kağan AKICI
Sayfa Adedi	102

Bu uzmanlık tezi, makine öğrenmesi ve yapay zeka teknolojilerinin siber güvenliğin sağlanmasındaki kullanımını ve bu teknolojilerin ülkemiz için sağladığı avantajları incelemektedir. Yapılan araştırmalar sonucunda, makine öğrenmesi ve yapay zeka teknolojilerinin siber güvenlik tehditlerini tespit etmek ve bu tehditlere karşı önlem almak için etkili araçlar olduğu ortaya konulmuştur. Bu teknolojiler, büyük veri setlerini analiz ederek anormallikleri ve potansiyel tehditleri hızlı bir şekilde tespit edebilme kapasitesine sahiptir. Sürekli öğrenen ve güncellenen sistemler sayesinde yeni ve gelişen tehditlere karşı proaktif bir savunma sağlanabilmektedir.

Tezde, Amerika Birleşik Devletleri (ABD), Birleşik Krallık, Fransa ve Türkiye'nin siber güvenlik uygulamaları araştırılmış ve makine öğrenmesi uygulamaları incelenmiştir. Ülkemiz için öneriler arasında eğitim ve farkındalık çalışmalarının artırılması, makine öğrenmesi ve siber güvenlik konularında uzmanlaşmış profesyonellerin yetiştirilmesi, kaliteli veri setlerine erişim sağlanması ve kamu-özel sektör iş birliğinin teşvik edilmesi bulunmaktadır. Ayrıca, yerli ve milli makine öğrenmesi çözümlerinin geliştirilmesi, bulut tabanlı çözümlerle büyük veri setlerinin işlenmesi ve depolanması önerilmiştir.

Bu öneriler, ülkemizin siber güvenlik alanındaki kapasitesini artırarak daha güvenli ve dirençli bir siber savunma yapısının oluşturulmasına katkı sağlayacaktır. Tez, makine öğrenmesi ve yapay zeka teknolojilerinin siber güvenlik alanında nasıl

kullanılabileceğini ve bu teknolojilerin ülkemiz için sağladığı potansiyel faydaları irdelemektedir.

ABSTRACT

BİLGİ TEKNOLOJİLERİ VE İLETİŞİM KURUMU	
Tezin Adı	Makine Öğrenmesi Kapsamında Yapay Zeka Teknolojilerinin Siber Güvenliğin Sağlanmasında Kullanımı, Ülke Uygulamaları ve Ülkemiz İçin Öneriler
Türü	Bilişim Uzmanlık Tezi
Yazar	Doğukan Ömer Gür
Teslim Tarihi	29.06.2024
Anahtar Kelimeler	Siber Güvenlik, Yapay Zeka, Makine Öğrenmesi
Tez danışmanı	Hasan Kağan AKICI
Sayfa Adedi	102

This thesis thoroughly examines the application of machine learning and artificial intelligence technologies in ensuring cybersecurity and the advantages these technologies bring to our country. The research demonstrates that machine learning and artificial intelligence technologies are effective tools for detecting cybersecurity threats and taking preventive measures against these threats. These technologies can quickly identify anomalies and potential threats by analyzing large datasets. Systems that continuously learn and update can provide proactive defense against new and evolving threats.

The thesis compares the cybersecurity practices of the USA, the UK, France, and Türkiye, evaluating these countries' machine learning applications. Recommendations for our country include increasing education and awareness efforts, training professionals specialized in machine learning and cybersecurity, providing access to quality datasets, and encouraging public-private sector collaboration. Additionally, it is suggested to develop domestic machine learning solutions, process and store large datasets using cloud-based solutions and engage in international research and development projects for knowledge exchange. Collaborations between universities and industry organizations are also encouraged to translate academic research into practical applications.

These recommendations aim to enhance our country's cybersecurity capacity, creating a more secure and resilient cyber defense infrastructure. The thesis highlights how

machine learning and artificial intelligence technologies can be utilized in the field of cybersecurity and the potential benefits these technologies offer for our country.

TEŞEKKÜR

Tez boyunca yanımda olan ve bu süreçte bana her türlü yardımcı olan Hasan Kağan Akıcı, Bahadır Cinoğlu ve Hasan Yılmaz'a bu noktaya gelmemi sağlayan annem ve halama, ve ilgili olduğum konularda çalışma imkanı sağlayan Demet Kabasakal ve Daire Başkanımız Mahmut Esat Yıldırım'a teşekkür ederim.

TABLOLARIN LİSTESİ

ŞEKİLLERİN LİSTESİ

KISALTMALARIN LİSTESİ

AE	Otokodlayıcı (*Autoencoder*)
APT	Gelişmiş Kalıcı Tehdit (*Advanced Persistent Threat*)
AI	Yapay Zeka (Artificial Intelligence)
AIS	Otomatik Gösterge Paylaşımı (*Automated Indicator Sharing*)
AS&F	Puanlama ve Geri Bildirim (*Assessment & Feedback*)
BT	Bilgi Teknolojileri
BTK	Bilgi Teknolojileri ve İletişim Kurumu
CNN	Evrişimli Sinir Ağları (*Convolutional Neural Networks*)
CISA	Siber Güvenlik ve Altyapı Güvenliği Ajansı (*Cybersecurity and Infrastructure Security Agency*)
CSAIL	Bilgisayar Bilimi ve Yapay Zeka Laboratuvarı (*Computer Science and Artificial Intelligence Laboratory*)
C2	Komuta ve Kontrol (*Command and Control*)
DBN	Derin İnanç Ağları (*Deep Belief Networks*)
DHS	Amerika Birleşik Devletleri İç Güvenlik Bakanlığı (*United States Department of Homeland Security*)
ENISA	Avrupa Siber Güvenlik Ajansı (*European Union Agency for Cybersecurity*)
LSTM	Uzun-Kısa Süreli Bellek (*Long Short-Term Memory*)
MIT	Massachusetts Teknoloji Enstitüsü (*Massachusetts Institute of Technology*)
MGK	Milli Güvenlik Kurulu
NATO	Kuzey Atlantik Antlaşması Örgütü (*North Atlantic Treaty Organization*)
NCPS	Ulusal Siber Güvenlik Koruma Sistemi (*National Cybersecurity Protection System*)
NCSC	Ulusal Siber Güvenlik Merkezi (*National Cyber Security Centre*)
NLP	Doğal Dil İşleme (*Natural Language Processing*)
NIST	Amerika Birleşik Devletleri Ulusal Standartlar ve Teknoloji Enstitüsü (*National Institute of Standards and Technology*)

NIDS	Ağ Saldırı Tespit Sistemleri (*Network Intrusion Detection Systems*)
NSA	Ulusal Güvenlik Ajansı (*National Security Agency*)
IDS	Saldırı Tespit Sistemi (*Intrusion Detection System*)
IPS	Saldırı Önleme Sistemi (*Intrusion Prevention System*)
ITU	Uluslararası Telekomünikasyon Birliği (*International Telecommunication Union*)
IoT	Nesnelerin İnterneti (*Internet of Things*)
PII	Kişisel Olarak Tanımlanabilir Bilgi (*Personally Identifiable Information*)
RNN	Yinelemeli Sinir Ağı (*Recurrent Neural Network*)
TDK	Türk Dil Kurumu
TFRE	Tehdit Odaklı Tersine Mühendislik (*Threat-Focused Reverse Engineering*)
SCADA	Gözetleyici Kontrol ve Veri Toplama Sistemi (*Supervisory Control and Data Acquisition*)
SOC	Güvenlik Operasyon Merkezi (*Security Operations Center*)
SSL	Güvenli Yuva Katmanı (*Secure Sockets Layer*)
VPN	Sanal Özel Ağ (*Virtual Private Network*)
XDR	Kapsamlı Algılama ve Yanıt (*Extended Detection and Response*)
YZ	Yapay Zeka

GİRİŞ

Siber güvenlik, günümüz dünyasında bilgi teknolojileri ve internetin yaygınlaşmasıyla birlikte büyük önem kazanmıştır. Bilgi teknolojileri ve iletişim ağlarının hızla gelişmesi, kişisel ve kurumsal bilgilerin dijital ortamda saklanması ve paylaşılması gereksinimini doğurmuştur. Bu durum, siber güvenlik tehditlerine karşı duyarlılığı artırmış ve bu tehditlere karşı etkili önlemler alınmasını zorunlu hale getirmiştir. Siber güvenlik, yalnızca teknik bir konu olmanın ötesinde, ulusal güvenlik, ekonomik istikrar ve bireylerin mahremiyetini koruma açısından da kritik bir rol oynamaktadır.

Yapay zeka ve makine öğrenmesi, siber güvenlik alanında devrim niteliğinde çözümler sunmaktadır. Yapay zeka, insan zekasını taklit eden ve belirli görevleri yerine getirmek için programlanabilen sistemlerdir. Makine öğrenmesi ise, bu sistemlerin deneyimlerinden öğrenerek performanslarını geliştirmelerini sağlayan bir yapay zeka dalıdır. Siber güvenlik alanında yapay zeka ve makine öğrenmesi uygulamaları, tehditlerin tespiti, saldırıların önlenmesi ve anomali tespiti gibi birçok alanda kullanılmaktadır.

Bu tezde, yapay zeka teknolojilerinin siber güvenliğin sağlanmasındaki rolü ve bu teknolojilerin ülkemiz için sağladığı potansiyel faydalar incelenmiştir. Tez kapsamında, Amerika Birleşik Devletleri, Birleşik Krallık, Fransa ve Türkiye'nin siber güvenlik uygulamaları ve makine öğrenmesi teknolojilerini nasıl kullandıkları araştırılmıştır. Ayrıca, ülkemiz için öneriler sunulmuş ve bu önerilerin siber güvenlik kapasitesini artırmada nasıl bir etki yaratabileceği değerlendirilmiştir.

Tezin amacı, siber güvenlik alanında yapay zeka teknolojilerinin kullanımı hakkında kapsamlı bir anlayış geliştirmek ve bu teknolojilerin ülkemiz için nasıl daha etkili bir şekilde kullanılabileceğine dair öneriler sunmaktır. Bu bağlamda, eğitim ve farkındalık çalışmalarının artırılması, makine öğrenmesi ve siber güvenlik konularında uzmanlaşmış profesyonellerin yetiştirilmesi, kaliteli veri setlerine erişim sağlanması ve kamu-özel sektör iş birliğinin teşvik edilmesi gibi konulara odaklanılmıştır.

Sonuç olarak, yapay zeka ve makine öğrenmesi teknolojilerinin siber güvenlik alanında etkin bir şekilde kullanılması, ülkemizin siber güvenlik kapasitesini artırarak daha güvenli ve dirençli bir siber savunma yapısının oluşturulmasına katkı sağlayacaktır. Bu tez, bu hedef doğrultusunda yapılan çalışmaları ve önerileri kapsamlı bir şekilde ele almaktadır.

1. SİBER GÜVENLİK VE YAPAY ZEKA

1.1. Siber Güvenlik

Bilgi teknolojileri (BT) ve internet hayatın her alanında kullanılmakta, sürekli bir şekilde gelişmekte ve günümüz toplumundaki yerini yeni uygulama ve yöntemler sayesinde sürekli derinleştirmektir. Uluslararası Telekomünikasyon Birliğinin (*International Telecommunications Union* - ITU) 2022 tahminine göre Dünya genelinde 5,3 milyar insan, toplam nüfusun yaklaşık yüzde altmışaltısı, internet kullanmaktadır ve bu sayı 2019 yılından bu yana yüzde 24 artmıştır(Sector, 2022). Geniş kitlelere yayılmış ve yaşam (Fischer, 2016)tarzımızı köklü bir şekilde değiştirmiş olan bu sistemler politika geliştiriciler için bir öncelik haline gelmiş bulunmaktadır. Kuzey Atlantik Antlaşması Örgütü (*North Atlantic Treaty Organization* - NATO) interneti "Ulusal altyapının hayati bir parçası ve toplumsal-ekonomik büyümenin ve gelişmenin sağlayıcısı olan kritik bir milli kaynak olarak tanımlamıştır"(Ziolkowski & NATO Cooperative Cyber Defence Centre of Excellence, 2012). Bu büyüklükte öneme sahip bir kaynağın kötü niyetli tarafların ya da devlet destekli oluşumların hedefine girmesi ise kaçınılmazdır. Ayrıca doğrudan bir grup ya da ülke tarafından hedef alınmasa dahi bu hizmetlerde aksamaya veya sızıntıya neden olabilecek zafiyetlerin tespit edilip düzeltilmesi gerekmektedir. Bu aktörlere ya da zafiyetlere karşı geçekleştirilen çalışmaların tamamı siber güvenlik olarak tanımlanabilir. Yapılan çalışmaların büyüklüğüne örnek olarak, son araştırmalara göre sektörün piyasa değerinin iki trilyon Amerikan Dolarına ulaşabileceği öngörülmektedir(Aiyer vd., 2022).

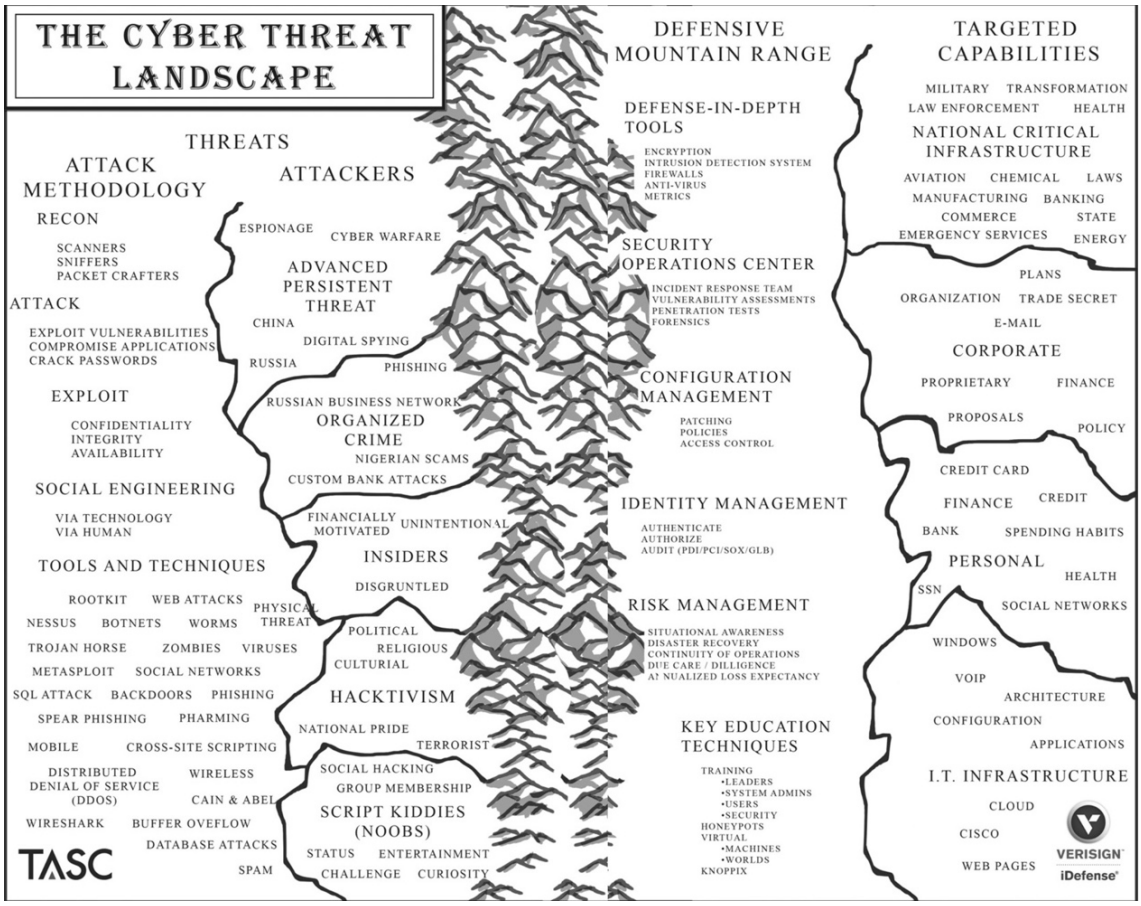
Siber güvenliğin pek çok tanımı olmasına karşın ITU tarafından 2008 yılında yayımlanan *Overview of cybersecurity* başlıklı raporda siber güvenlik aşağıdaki gibi tanımlanmıştır;

> "*Siber güvenlik, siber ortamı ve kuruluş ile kullanıcı varlıklarını korumak için kullanılabilecek araçlar, politikalar, güvenlik kavramları, güvenlik önlemleri, kılavuzlar, risk yönetimi yaklaşımları, eylemler, eğitim, en iyi*

> *uygulamalar, teminat ve teknolojiler bütünüdür. Kurum ve kullanıcı varlıkları, bağlı bilgi işlem cihazlarını, personeli, altyapıyı, uygulamaları, hizmetleri, telekomünikasyon sistemlerini ve siber ortamda iletilen ve/veya depolanan bilgilerin bütününü içerir. Siber güvenlik, siber ortamdaki ilgili güvenlik risklerine karşı kuruluşun ve kullanıcının varlıklarının güvenlik özelliklerinin elde edilmesini ve sürdürülmesini sağlamak için gayret gösterir. Genel güvenlik hedefleri aşağıdakilerden oluşur:*
> *- Erişilebilirlik*
> *- Bütünlük, özgünlük ve reddedilmeme özelliklerini içerebilir*
> *- Gizlilik.*"(Study Group, 2008)

Kamuoyunda siber güvenlik kavramı, mahremiyet, bilgi paylaşımı, istihbarat toplama ve izleme gibi kavramlarla sıklıkla karıştırılmaktadır. Siber güvenlik araçları mahremiyetin korunması konusunda önemli rol oynarken, kimi zaman mahrem kabul edilebilecek bilgilerin siber güvenliğin sağlanması amacıyla kullanılabileceği göz ardı edilmemelidir. Benzer şekilde izleme ve istihbarat faaliyetleri bazı durumlarda siber güvenlik uzmanlarının teknik bilgisine başvursa da bu kavramlar siber güvenlik için bir araç olarak olup siber güvenliğin tanımının bir parçası değildir.

Şekil 1.1. Siber Tehdit Manzarası



Kaynak: (Andress & Winterfeld, 2014)

1.1.1. Siber güvenlik tehdit aktörleri ve tehditleri

Siber güvenlik tehditleri hakkında ulusal ve uluslararası kuruluşlar tarafından çeşitli tanımlamalar yapılmıştır. Amerika Birleşik Devletleri Ulusal Standartlar ve Teknoloji Enstitüsü (National Institute of Standards and Technology - NIST) tehdidi; "*Yetkisiz erişim, imha, ifşa, bilginin değiştirilmesi ve/veya hizmet reddi yoluyla bir bilgi sistemi üzerinden kurumsal operasyonları (misyon, işlevler, imaj veya itibar dahil), kurumsal varlıkları veya bireyleri olumsuz etkileme potansiyeline sahip herhangi bir durum veya olay. Ayrıca, bir tehdit kaynağının belirli bir bilgi sistemi güvenlik açığından başarılı bir şekilde yararlanma potansiyeli.*(Computer Security Division, 2006)" olarak tanımlamıştır. Bu tehditlerin gelebileceği aktörler motivasyonlarına ya da ne kadar sofistike olduklarına göre kategorize edilebilir. Siber tehdit aktörleri

bilgisayarların işlem gücünü kullanmak, bilgi sızdırmak ve/veya manipüle etmek ya da casusluk ve şantaj gibi farklı amaçlar doğrultusunda eylemler gerçekleştirmektedir. Tehdit aktörleri doğrudan bir kurum ya da kişiyi hedef alabilmekle birlikte birçoğu fırsatçı bir yaklaşımla savunmasız sistemleri hedef almaktadır.

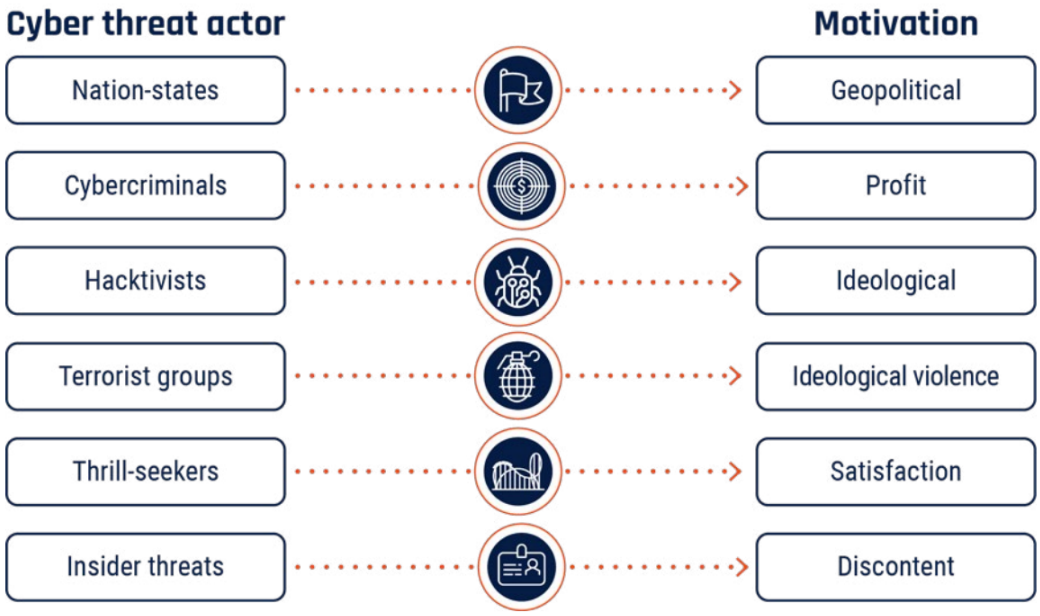
Ulaştırma ve Altyapı Bakanlığı tarafından hazırlanan Ulusal Siber Güvenlik Stratejisi ve Eylem Planı'nda (2020-2023) ise, siber tehdit "Bir kurumun veya sistemin zarar görmesi ile sonuçlanabilecek istenmeyen bir siber olayın potansiyel nedeni." olarak tanımlanmıştır(UAB, 2020).

Siber tehdit aktörleri yetenek ve gelişmişlik açısında geniş bir spektruma yayılmışlardır. Söz konusu aktörler tek başlarına, küçük ekiplerle ya da büyük organizasyonlara bağlı olarak faaliyet gösterebilmektedir. Bazı durumlarda ileri derecede sofistike aktörler açık kaynaklı ya da kolay erişilebilir araçlar tercih etmektedir. Bu araçlar çoğu durumda yeterince etkili olmakla birlikte daha da önemlisi siber saldırıyı gerçekleştirenlere karşı isnadı zorlaştırmaktadırlar.

Siber tehdit aktörlerinin başında Gelişmiş Kalıcı Tehdit (*Advanced Persistent Threat* - APT) olarak nitelendirilen gruplar yer almaktadır. En sofistike yöntem ve imkanlara sahip olan bu grupların öne çıkan özellikleri; amaçlarına ulaşmak konusunda ısrarcı olmaları, uzun vadeli operasyonlar düzenlemeleri, hedeflerinin güvenlik önlemlerine göre yeni yöntemler geliştirip, uyum sağlamaları ve hedeflerine ulaşana kadar saldırdıkları sistemde kalıcı olmaları olarak sıralanmaktadır. APT grupları devletlere, organizasyonlara ya da örgütlere bağlı olarak ya da bağımsız bir şekilde çalışmalarını sürdürmektedir(Communications Security Establishment (Canada), 2022; Computer Security Division, 2011). Bir diğer siber tehdit unsuru ise Devlet destekli siber tehdit aktörleridir, bu aktörler doğrudan bir kurum ya da kuruluşa bağlı olarak çalışmakta olup bağlı oldukları devletin imkanları ile orantılı olarak çok sofistike olabilmektedir. Bu aktörler bağlı oldukları devletin çıkarları doğrultusunda çalışmalar yürütürler. Mali motivasyonun en ön planda olduğu siber tehdit aktörleri ise siber suçlulardır. Bu aktörler sofistikelik açısından farklı seviyelerde olabilirler ve dolandırıcılık ve/veya şantaj gibi doğrudan maddi çıkar sağlamak üzere operasyonlar yürütmektedirler.

Bilgisayar korsanı (*hacker*) ve eylemci (*activist*) kelimelerinin birleşmesi ile oluşturulan *hacktivist* kelimesi ise ideolojik motivasyonlarla hareket eden ve genellikle hedeflerine maddi hasardan çok itibar hasarı yaşatmaya çalışan grupları tanımlamak için kullanılmaktadır(Andress & Winterfeld, 2014). Bu aktörlerin yanı sıra sıklıkla karşılaşılan bir diğer siber tehdit aktörü ise *insider* olarak tanımlanan ve saldırıya maruz kalan yapının içerisinde çalışan kişilerdir. Bu aktörler kritik bilgilere kolay erişebilir ya da başka birinin erişebilmesi için içerden açık kapı bırakabilmektedirler. Maddi ve manevi memnuniyetsizlik bu aktörlerin motivasyon kaynaklarının başında gelmektedir(Andress & Winterfeld, 2014).

Şekil 1.2. Siber Tehdit Aktörleri ve Motivasyonları



Kaynak: (Communications Security Establishment (Canada), 2022)

Siber tehdit, bir sistemin veya içerdiği bilgilerin kullanılabilirliğini, bütünlüğünü veya gizliliğini değiştirerek bir bilgi sisteminin güvenliğini tehlikeye atmayı veya genel olarak dijital yaşamı kesintiye uğratmayı amaçlayan faaliyetlerdir. Bu tehditler sürekli değişmekte ve gelişmekte olup 2022 yılı için Avrupa Siber Güvenlik Ajansı (*European Union Agency for Cybersecurity* - ENISA) tarafından başlıca sekiz tehdit tanımlanmıştır.

Bunlar;

- Fidye yazılımları (*Ransomware*); bu tür yazılımlar kullanıcıların kendi dosyalarına, işletim sistemlerine veya cihazlarına erişimini engellemek veya kısıtlamak için tasarlanmış ve erişimin yeniden kazanılması için fidye ödenmesini talep eden kötü amaçlı bir yazılım türüdür(Paquet-Clouston vd., 2019). Genel olarak fidye yazılımları, temel bilgisayar işlevleri için gerekli olan dosyaları şifreleyen kilitleyici (*locker*) fidye yazılımları ve kullanıcının hassas dosyalarını şifreleyen kriptografik fidye yazılımları olarak iki geniş kategori altında sınıflandırılabilmektedir (Razaulla vd., 2023). Fidye yazılımları ilk kez 1989 yılında ortaya çıkmış ve günümüzde de varlığını güçlü bir şekilde sürdürmektedir. Şifreleme yöntemleri, hızlı yayılma, tespitten kaçma ve kurbanları fidye ödemeye zorlama yetenekleri bakımından giderek gelişmektedirler (Kharraz Amin vd., 2015). Ortaya çıkışlarından bu yana fidye yazılımları siber tehditler arasında en önemli ve yaygın yöntemlerden biri olarak yerini korumaktadır.
- Zararlı yazılımlar (*Malware*); kötü niyetli/zararlı (*Malicious*) ve yazılım (*software*) kelimelerinin birleştirilmesi ile oluşturulan *malware* kelimesi veri çalmak, bilgisayarlara ve sistemlere zarar vermek ya da doğrudan kişilere zarar vermek için özel olarak hazırlanan yazılımların tamamına verilen isimdir. Kötü amaçlı yazılımlar internetin ilk yaygınlaştığı günlerden itibaren ortaya çıkmışlardır fakat ilk çıktıkları dönemde son kullanıcılar nadiren hedef olarak seçilmiştir. Son kullanıcılar yerine, bankacılık, finans ve devlet gibi sektörler, bu yazılımların ana hedefi olmuşlardır, ancak, kötü amaçlı yazılım ekosistemi zaman içinde değişerek hedef alanını genişletmiştir. İlk ortaya çıktığında bu yazılımların hedefinde çoğunlukla para varken şu an verinin artan değeri sayesinde zararlı yazılımların birincil hedefi veri halini almıştır(Alrawili vd., 2022).
- Sosyal mühendislik; siber güvenlik bağlamında kullanıldığı zaman sosyal mühendislik, kötü niyetli amaçlarla kullanılmak üzere, bir ya da birçok kişiyi gerçek olmayan bir şeye inandırarak kişisel bilgilerine erişilmesi süreci olarak tanımlanabilir. Bu süreç genellikle, kullanıcıların belgeleri, dosyaları veya e-postaları açmaları, web sitelerini ziyaret etmeleri veya yetkisiz kişilere

sistemlere veya hizmetlere erişim izni vermeleridir. Bu yöntemler teknolojinin kötüye kullanımı olarak görünseler de başarılı olmaları için her zaman insan kusuru bulması gerekmektedir(Hatfield, 2018).

- Veriye karşı tehditler; verilere yetkisiz erişim ve verilerin ifşa edilmesinin yanı sıra sistemlerin davranışlarına müdahale etmek için verilerin manipüle edilmesi amacıyla veri kaynaklarını hedef alan tehditlerin tamamını oluşturmaktadır. Listede yer verilen tehditlerin çoğunluğunun ya sonucu ya da sebebi olabilen veri tehditleri genel olarak veri ihlali ya da veri sızıntısı olarak karşımıza çıkmaktadır. Veri ihlali, kötü niyetli aktör tarafından yetkisiz erişim elde etmek suretiyle hassas ve/veya gizli verilerin ele geçirilmesi ya da ifşa edilmesi ile gerçekleşen kasıtlı bir saldırıdır. Veri sızıntısı ise yanlış prosedürler, güvenlik açıkları veya insan hataları nedeniyle hassas ve/veya gizli verilerin kasıtsız olarak yayınlanması için kullanılmaktadır(Lella vd., 2022).
- Erişilebilirliğe karşı tehditler-hizmet reddi (Denial of Service-DoS); Hizmet reddi saldırıları, bir saldırganın bir hizmetin meşru kullanıcılarının o hizmetin kaynaklarını kullanmasını engellemeye yönelik olarak gerçekleştirilen saldırılardır(Needham, 1993). Bunlara örnek olarak aşağıdakiler sayılabilir;
    - Bir ağı aşırı yükleyerek meşru ağ trafiğini engelleme girişimleri.
    - İki makine arasındaki bağlantıların bozulmaya çalışılması yoluyla hizmet engellenmesi.
    - Belirli bir kişinin bir hizmete erişimini engelleme girişimleri.
    - Belirli bir sisteme veya kişiye verilen hizmeti kesintiye uğratma girişimleri(Lau vd., 2000).
- İnternetin erişilebilirliğine karşı tehditler, hizmet reddi saldırıları dışında kalan diğer erişilebilirlik saldırıları ENISA tarafından internetin erişilebilirliğine karşı tehdit başlığı altında değerlendirilmektedir.
- Dezenformasyon, bilinçli bir şekilde insanları kandırmak ya da yanıltmak amacıyla yayınlanan çoğu zaman gerçek gözükecek şekilde hazırlanan yanlış bilgilerin yayılmasıdır(Pacepa & Rychlak, 2003). Dezenformasyon başta iletişim sorunu gibi görünmekle birlikte günümüzde insanların bilgi kaynağı

olarak kullandığı platformlar göz önüne alındığında bir siber güvenlik tehditti olarak da değerlendirilebilmektedir. Sosyal medya ve internet sitelerinin çalışma prensibinin tıklama ve görüntülenme kazanılması üzerine olması, çoğu zaman daha sansasyonel ve ilgi çekici olan dezenformasyonların hızla yayılmasına olanak sağlamaktadır.

- Tedarik zinciri saldırıları, tedarik zinciri bir ürün ya da hizmetin üreticisinden müşterisine kadar giden bütün süreç ve aktörleri ifade etmektedir. Siber güvenlik düşünüldüğünde yazılım ve donanımlar, bulut ya da yerel depolama alanları ürün dağıtımı yapan çevrimiçi mağazalar, web uygulamaları ya da yönetim yazılımları tedarik zincirinin parçasıdır. Tedarikçi, tedarikçi varlıkları, müşteri ve müşteri varlıkları, tedarik zincirlerinin dört temel öğesi olarak değerlendirilmektedir. Bir saldırının tedarik zinciri saldırısı sayılabilmesi için en az iki öğenin etkilenmesi gerekmektedir. Hem tedarikçiye hem de müşteriye zarar veren bu saldırılar, kurulan tedarik zinciri sayesinde hızla yayılarak çok tehlikeli bir boyut kazanabilmektedir (ENISA, 2021).

ENISA tarafından tanımlanan bu tehditlerin siber uzayda yer alan tüm tehditlere karşılık gelmediği ve bu alanın sürekli olarak değiştiği unutulmamalıdır.

1.1.2. Siber güvenlik konseptinin gelişimi

Siber tehdit ortamı değiştikçe, siber güvenlik konsepti de hızla değişim göstermektedir. Bilgi güvenliği konseptinin 1980'li yıllardan günümüze kadar değişiminin incelenmesinde öne çıkan bir görüş, değişimin beş dalga olarak incelenmesidir(Joe Turner vd., 2010).

İlk dalga 1980'li yıllarda ortaya çıkmıştır ve *Technical Wave* – Teknik Dalga olarak adlandırılmaktadır. İlk dalga tamamen kullanılan bilgisayar ve cihazların korunmasına odaklanmaktadır. Bilgi güvenliği söz konusu olduğunda ana makinelere giriş yapılması için kimlik doğrulama ve ilkel yetkilendirme işlemleri gibi teknik çözümler ön plana çıkmaktadır. Bu önlemler ise bilgisayarları kullanan teknik personel ile sınırlı kalarak kapsayıcı olamadığı değerlendirilmektedir. Bu dönemde, politika, prosedür ve

farkındalık gibi konular ise gereken önemi görmemektedir(S. H. (Basie) von Solms, 2010).

İkinci dalga ise 1990'ların ortasına kadar devam eden *Managmet Wave* – Yönetim Dalgasıdır. Dağıtık bilgi işlemin ve kişisel bilgisayarın gelişimi, bilgi güvenliği alanına pek çok dönüşümün yaşanmasına kapı açmıştır. Bilginin tek bir merkezde ya da iyi korunan bilgisayarda saklanmayıp ağlarla birbirine bağlı çok sayıda bilgisayara dağıtılması, daha önce var olmayan ciddi güvenlik riskleri yaratmıştır. Bilgi güvenliği, yönetimin dikkatini çeken bir konu haline gelmiş ve bilgi güvenliği yöneticileri ortaya çıkmıştır. Bu yeni kadrolar bilgi güvenliği politikaları ve prosedürleri oluşturmaya ve bilgi güvenliği departmanları da bulunduran organizasyon yapıları oluşturmuşlardır. Kurumlardaki bilgi güvenliğinin durumu hakkında raporlama yapma ihtiyacı da ortaya çıkmıştır. Bu gelişmeler elbette genel olarak bilgi güvenliğinin gelişmesine sebep olmuş ve bilgi güvenliğinin önemli bir yönetim boyutu sahip olduğunu ve güvenli bir ortam yaratmak için bu boyuttan tam olarak yararlanılması gerektiğini ortaya koymuştur(Joe Turner vd., 2010).

İkinci Dalga sırasındaki, bu gelişmeler nedeniyle sektör bilgi güvenliği alanında en iyi uygulamalar ve standardizasyon ile ilgili konuları araştırmaya başlamışlardır. Bu araştırmalar sonucunda kurumlar bilgi güvenliği planlaması yapmak adına sorular oluşturmaya başladılar(Rannenberg vd., 2010);

- Rakiplerimizle güvenlik açısından nasıl kıyaslanabiliriz?
- Bir bilgi güvenliği politikasında neler olmalıdır?
- Kurumun bilgi güvenliği durumunu nasıl bir resmi sertifika ile kanıtlayabiliriz?

Gibi sorular ile kurumsallaşmaya giden üçüncü aşamanın temelleri atılmıştır.
Bu süreçte, sistemin son kullanıcısı olan çalışanların rolü mercek altına alınarak ve bilgi güvenliğinin insan boyutunun önemi kabul edilmiştir.

Üçüncü dalga ise *Institutionalization*-Kurumsallaşma olarak karşımıza çıkmaktadır. Bilgi güvenliğinin sadece teknik bir konu olmadığı, çok boyutlu olduğu anlaşıldıkça, bir kurumun geleceğinin, sağlıklı büyümesinin ve stratejisinin bir parçası olarak Bilgi

güvenliğinin değerlendirilmesi, kurumsallaştırılarak kurumun kültür ve düşünce tarzının bir haline getirilmesi için çalışmalar başlamıştır. Bilgisiz çalışanların bilgi güvenliğini tehlikeye atabileceğinin anlaşılması ile bu süreçte tüm çalışanların eğitilmesi ve farkındalık arttırma çalışmaları gibi konseptler ortaya çıkarak yaygınlaşmıştır. Yine bu süreç içerisinde kurumların genel güvenlik seviyelerini belirlemek ve bunları raporlamak için yöntemler geliştirilmeye başlanmıştır. Bu yaşanan gelişmeler kurumsal yapılanmada bilgi güvenliğinin rolünü güçlendirmiştir. Gelişmelere paralel olarak ortaya çıkan yasal düzenlemeler ve standartlar bir sonraki aşama olan bilgi güvenliği yönetiminin oluşmasında rol oynamıştır(S. H. (Basie) von Solms, 2010).

Dördüncü dalgada ise bilgi teknolojilerine dair risk yönetimi ön plana çıkarak *Information Security Governance*-Bilgi Güvenliği Yönetimi konseptini oluşturmuştur. Bu dönemde bilgi güvenliği ile ilgili standart ve en iyi yöntemler yaygınlık kazanmıştır. Finans sektörünün hızla dijitalleşmesi, bilgi güvenliği kaynaklı ciddi maddi kayıplar yaşanması olasılığını da arttırmıştır. Üst yönetimler artık kurumları tarafından kullanılan bilgi teknolojilerinin güvenliğinin kendi sorumlulukları olduğunu kabul ederek kurum içi yönetimsel önlemleri devreye sokmuşlardır. Bilgi güvenliği böylece yönetimin kademesinin bir parçası olarak yer bulmuştur(B. von Solms, 2006).

Şimdiye kadar ele alınan dört dalganın ortak noktalarını inceleyecek olursak her zaman iç güvenliğin ön planda olduğu, hedeflenen kurumun sistemlerinin korunmasını amaçladığı ve sorumlunun kurum ve çalışanları olduğu anlaşılmaktadır. Fakat internetin hızla yayılması ve kurumların internet üzerinden hizmetlerini sağlamasının başlaması ile Siber çağın başladığı söylenmektedir. Bu dönüşüm bilgisayar korsanlarının iç güvenlik önlemlerini arttıran kurumlar yerine sayıları milyonları bulan son kullanıcılara yönelmesine sebep olmuştur. Farkındalığın çok düşük olduğu son kullanıcılar arasında özellikle sosyal mühendisliğin de etkisiyle siber suçlar hızla yayılmaktaydı. Bilgisayar korsanları arasında yayılan güçlü güvenliğe sahip kurumlar yerine deneyimsiz son kullanıcıya saldırılması, bilgi güvenliğinde yeni bir çağ başlatarak siber güvenlik kavramını ortaya çıkartmıştır.

Böylece beşinci ve son dalga olan *Cyber Security*-Siber Güvenlik dönemine gelinmiştir. Artık kurumların sadece kendilerini değil internet üzerinden hizmet verdikleri milyonlarca son kullanıcıyı da koruması gerekmektedir. Siber çağın getirdiği bir diğer yenilik ise suçlular arasında yeni yöntem ve araçların hızla yayılmasına olanak sağlaması ve bu suçluların bilgi paylaşımında bulunabileceği bir platform oluşturmasıdır. Yeni nesil bilgi güvenliği uzmanlarına düşen görev ise bu değişikliklere adapte olunmasıdır(S. H. (Basie) von Solms, 2010)

1.2. Yapay Zeka

İnsanoğlu diğer canlılardan kendini ayırırken öncelikle zekasına dikkat çekmektedir. Taksonomi biliminde kendini *homo sapiens* – bilge insan olarak adlandırması buna bir kanıt olarak gösterilebilmektedir. Zeka dilimize Arapçadan geçmiş bir kelime olup Türk Dil Kurumu tarafında hazırlanan Güncel Türkçe Sözlükte "*İnsanın düşünme, akıl yürütme, öğrenme, kavramları ve nesneleri zihinde canlandırabilme, objektif gerçekleri algılama, yargılama, sonuç çıkarma, bedeni kontrol edebilme, duyguları doğru algılayabilme, değerlendirebilme, icat edebilme vb. yeteneklerinin ve becerilerinin tamamı; anlak, dirayet, feraset"* olarak tanımlanmıştır(TDK, 2023).

Yapay zeka ise, insan beynini taklit etmek ve gerçek dünya problemlerini bütünsel bir insan yaklaşımıyla bilgisayarlara çözdürmeyi araştırmak için bir kavram olarak ortaya atılmıştır. Bu doğrultuda, bilgi teknolojileri ve fizyolojik zekanın bir araya getirilmeye çalışılması olarak düşünülebilmektedir. Yapay zeka uygulamaları günümüzde, büyük miktarda depolanmış verinin akıllıca kullanılmasını ve işlenmesini mümkün kılarak işlevsel araçların oluşturulmasını sağlamaktadır. Ayrıca, savunma, sağlık ve uzay araştırmaları gibi çeşitli alanlarda sorun çözmek için kullanılmaktadır.

Yapay zeka araştırmaları çok çeşitli olsa da günümüzde iki farklı ve yaygın türü gelişmiştir. Bir tarafta Massachusetts Teknoloji Enstitüsü ile bağdaştırılan; akıllı davranış sergileyen herhangi bir sistemi yapay zeka örneği olarak görülmesidir. Bu düşünce ekolüne göre, geliştirilen sistemin görevini insanlarla aynı şekilde yerine getirip getirmediği önemli değildir. Tek ölçüt, sistemin doğru sonuç vermesidir.

Özellikle; elektrik mühendisliği, robotik ve ilgili alanlardaki yapay zeka projelerinin sonuçlarında doğru sonuç ve tatmin edici performans hedeflenir, yapay zekanın nasıl çalıştığı ikinci plandadır. Bu yaklaşım zayıf yapay zeka olarak adlandırılmaktadır. Bir başka düşünce ekolü ise Carnegie-Mellon Üniversitesi'nin yapay zekaya yaklaşımıyla temsil edilir, bu yaklaşım ise sonuçla olduğu kadar biyolojik akla yatkınlıkla da ilgilenmektedir. Bu ekole göre, bir sistem akıllı davranış sergilediğinde, süreci insanlar tarafından kullanılan metodolojilere dayanmalıdır. Örneğin, işitme yeteneğine sahip bir sistem yapılması planlandığında. Zayıf yapay zekaya göre yalnızca sistemin performansı önemliyken, güçlü yapay zekaya göre işitme kanalı, kulak zarı ve kulağın diğer parçalarına eşdeğer olan ve her biri sistemde gerekli görevleri yerine getiren insan işitme sisteminin taklit edilmesi başarı kriteri olacaktır. Zayıf yapay savunucuları, inşa ettikleri sistemlerin başarısını yalnızca performanslarına göre ölçerken, güçlü yapay zeka savunucuları inşa ettikleri sistemlerin yapısıyla ilgilenmektedir(Lucci vd., 2022). Makine öğrenmesi daha sonuç odaklı bir yaklaşım olup bahse konu yaklaşımlardan zayıf yapay zeka kategorisinde olduğu değerlendirilebilir.

1.2.1. Makine öğrenmesi

Makine öğrenmesi bilgisayarların özel olarak programlamaya ihtiyaç duymadan tecrübeye dayalı olarak öğrenerek bir sorunu çözmesine odaklanan yapay zeka dalıdır(Rostamizadeh & Talwalkar, 2012). Makine öğrenmesi için çok çeşitli senaryolar bulunmakla birlikte bunları geniş olarak iki çeşide ayırabiliriz; gözetimli öğrenme ve gözetimsiz öğrenme. Öğrenme senaryoları bunlarla kısıtlı olmamakla birlikte anomali tespiti[1] gibi bazı alanlar hem gözetimli hem de gözetimsiz öğrenme yöntemlerini barındırabilirler(Rostamizadeh & Talwalkar, 2012).

Gözetimli öğrenmede eğitilen Makine öğrenmesi algoritması (model) önceden etiketlenmiş verileri kullanmaktadır. Model bu etiketli verilerden kendini tekrar eden

[1]Olağan, standart veya beklenenden sapan gözlemlerin, olayların veya veri noktalarının tanımlanmasıdır. Bu gözlemler veri kümesinin geri kalanıyla tutarsız olmaktadır.

yinelemeli bir süreçle öğrenir ve daha sonra yeni daha önce görmediği veriler üzerinde tahminler gerçekleştirir. Bu yöntem regresyon ve sınıflandırma sorunlarının çözümünde sıklıkla kullanılmaktadır(Rostamizadeh & Talwalkar, 2012).

Regresyon kullanılması demek, model tarafından tahmin edilecek çıktının değişkeninin sayısal bir değer olması durumudur. Örneğin, Ankara'daki konut fiyatlarının tahmin edilmesi veya belirli bir mahallede tüketilen enerji miktarının tahmin edilmesi gibi. Doğrusal regresyonun amacı sağlanan etiketli veri setinin kullanılarak kurulan modelin ne kadar iyi çalıştığına dair bir gösterge olan maliyet fonksiyonunu minimuma çekecek parametrelerin bulunmasıdır. Regresyon tek ya da çok değişkenli olabilmektedir. Tek değişkenli doğrusal regresyon $x(i)$ şeklinde oluşturulan denklemleri ifade etmektedir tek boyutlu bir vektör kullanılarak çözüme ulaşılmaya çalışılmaktadır. Bu tür bir model için kullanılacak veri setine örnek olarak insanların yaş ve boyu olan bir tablo kullanılabilir.

Tablo 1.1. Yaş Boy Tablosu

	Yaş	**Boy**
1	14	161
2	56	172
3	34	173
4	81	181
5	29	155

Yukarıda kullanılan veri setine doğrusal regresyon uygularsak i yaşı $x(i)$ fonksiyonunun sonucu ise boyu temsil etmektedir. Bu parametreleri kullanarak insanların yaşından boyunu tespit etmeye çalışan bir yapay zeka, daha özel olarak bir makine öğrenmesi algoritması eğitilebilmektedir.

Makine öğrenmesinde hipotez, girdi verilerini çıktı tahminlerine dönüştüren matematiksel bir fonksiyon veya modeldir. Veri seti kullanılarak yeni değerler oluşturulabileceğini temsil eder. Hipotez tipik olarak modelin davranışını karakterize eden bir parametreler topluluğu olarak ifade edilmektedir. Parametreler ise, modelin

içinde bulunan ve değeri önceden verdiğimiz verilerden tahmin edilebilen değişkenlerdir. Parametreler makine öğrenmesi algoritmalarının ana unsurlarıdır. Asıl olarak parametreler modelin geçmiş eğitim verilerinden öğrenilen kısmıdır. Makine öğrenmesi literatüründe modeli hipotez, parametreleri ise hipotezin belirli bir veri kümesine uyarlanması olarak düşünebiliriz. Maliyet fonksiyonu, bir veri seti için bir makine öğrenmesi modelinin performansını ölçmektedir. Maliyet fonksiyonu, tahmin edilen ve beklenen değerler arasındaki hatayı ölçer ve bu hatayı tek bir gerçek sayı şeklinde sunar. Probleme bağlı olarak, maliyet fonksiyonu birçok farklı şekilde oluşturulabilmektedir.

Bu bilgiler göz önüne alındığında yaş verisinden boy tahmini yapan modelimiz aşağıdaki şekilde ifade edilmektedir;

Hipotez; $h_\theta(x) = \theta_0 + \theta_1 x$

Parametreler; θ_0, θ_1

Maliyet Fonksiyonu; $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2$

Amaç; $\min_{\theta_0,\theta_1} J(\theta_0, \theta_1)$

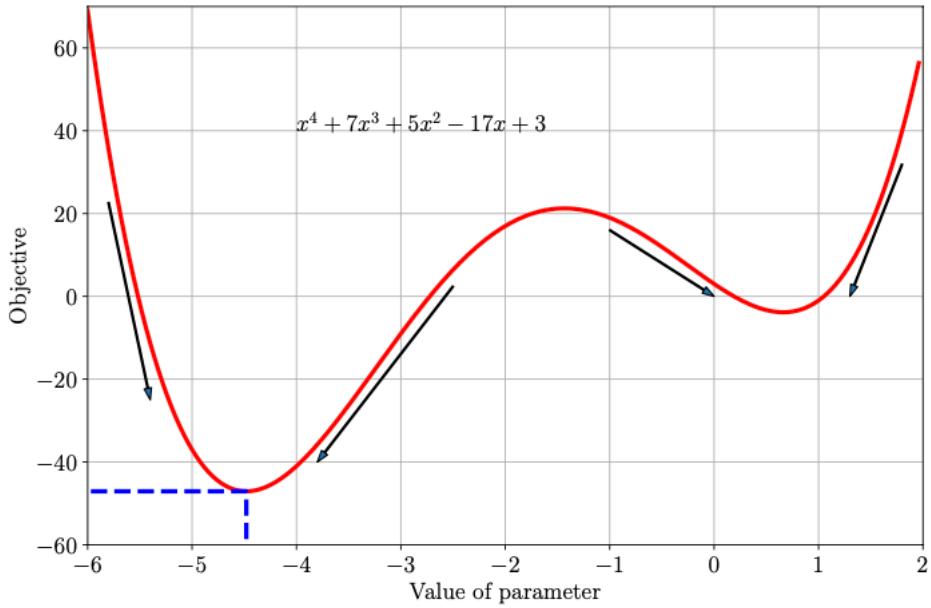
Bu fonksiyonlar göz önüne alındığında karşılaşılan sorun ise θ (teta) parametre değerlerinin nasıl bulunacağıdır. Gradyan Azalması, farklı türde fonksiyonları minimize etmek için kullanılan genel bir algoritmadır. Doğrusal regresyonda, model parametrelerini otomatik ve eş zamanlı olarak güncellemek için bir gradyan azalış algoritması kullanılır. Bu algoritma yinelemeli bir algoritmadır; her yineleme için maliyet fonksiyonunun türevi değerlendirilir ve model parametreleri güncellenir. Bu yeniden değerlendirme, maliyet fonksiyonunun varsa global optimal noktası bulunana kadar tekrarlanır(Deisenroth vd., 2020).

$$Minimum\ noktaya\ yakınsama\ olana\ kadar\ tekrar\ et;$$

$$\left\{\begin{array}{l} \theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) \\ \theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right).x^{(i)} \end{array}\right\}$$

Gradyan azalma algoritmasını oluşturan iki temel unsur öğrenme oranı/alfa (α) ve maliyet fonksiyonunun kısmi türevleridir. Öğrenme oranı, model parametrelerini güncellerken ne kadarlık bir adım boyutu alacağımızı belirler. Yukarıdaki toplama terimi ile gösterilen kısmi türev ise en dik eğimin yönünü verir.

Şekil 1.3. Gradyan Azalmada Yönler



Kaynak: (Deisenroth vd., 2020)

α çok küçük seçilirse, yerel bir minimuma yakınsamak çok zaman alabilir, yapmamız gereken işlem sayısı artar ve daha çok adımda işlemi tamamlamamız gerekir, bunun sonucunda gradyan azalma yavaş olur. α çok büyük seçilirse, optimum hedefi aşabilir ve sapabiliriz, minimum değerin üzerinden atlayabilir ve asla yakınsamayabiliriz.

Tahmin için kullanacağımız eğitim verilerimizde tek bir özellik yerine birden daha fazlası olduğunda, verilerin her biri artık tek bir değer yerine bir vektör haline gelir.

Bu gibi durumlarda her bir veri n boyutlu bir özellik vektörü haline gelir; burada n = eğitim setindeki özellik sayısıdır. Böyle durumlar çok değişkenli Doğrusal Regresyon olarak tanımlanmaktadır. Aşağıda, her bir eğitim verisi için n özellik sayısının dört olduğu ve tahminin maaş olduğu örnek bir veri kümesi bulunmaktadır.

Tablo 1.2. Maaş Tahmin Tablosu

	Cinsiyet	Sınav Puanı	Tecrübe Yılı	Birim	Maaş
1	K	89	2	ARGE	65
2	K	73	5	Pazarlama	54
3	E	66	9	Pazarlama	55
4	K	95	14	ARGE	78
5	E	88	2	Üretim	82

Çok değişkenli doğrusal regresyon için hipotez, aşağıda gösterilen çok değişkenli doğrusal regresyonun genişletilmiş halidir, n özellik sayısını ifade eder;

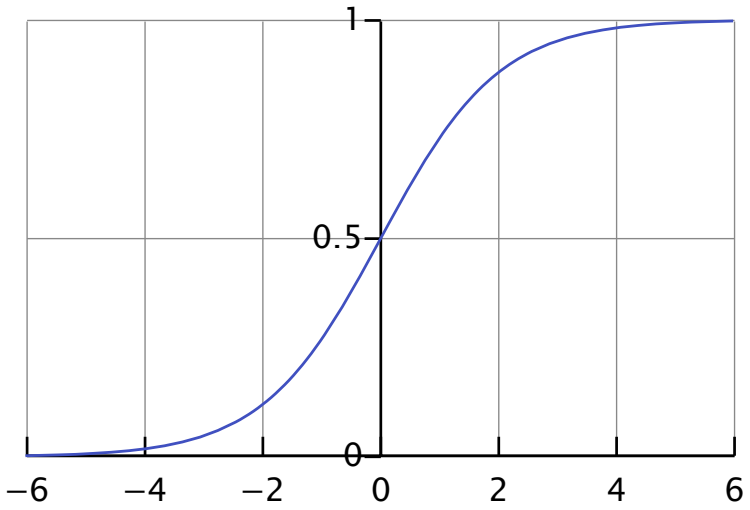
Hipotez; $h_\theta(x) = \theta^T x = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$

Parametreler; $\theta_0, \theta_1, \theta_2, \ldots, \theta_n$

Maliyet Fonksiyonu; $J(\theta_0, \theta_1, \ldots, \theta_n) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2$

Makine öğrenmesinin önceki verilere dayanarak tahminde bulunması dışında en çok kullanıldığı bir diğer alan ise sınıflandırma işlemleridir. Sınıflandırmada amaç, modele verdiğimiz girdinin, önceden belirlediğimiz sınıflardan hangisine ait olduğunu belirlemektir. Örneğin, bir görüntünün köpek mi kedi mi, kırmızı mı beyaz mı, bir e-postanın spam mı yoksa gerçek mi ya da bir hastada kanser var mı yok mu gibi sorulara cevap bulmak için kullanılmaktadır(Rostamizadeh & Talwalkar, 2012). Makine öğrenmesi alanında sınıflandırma sorularının çözümü için sıklıkla kullanılan bir modelse lojistik regresyon modelidir. Lojistik regresyon sonuç değerleri 0 ile 1 arasına sıkıştırılmış sigmoid veya lojistik fonksiyona dayanan bir sınıflandırma algoritmasıdır.

Şekil 1.4. Lojistik regresyon grafiği



Hipotez; $h_\theta(x) = g(\theta^T x)$

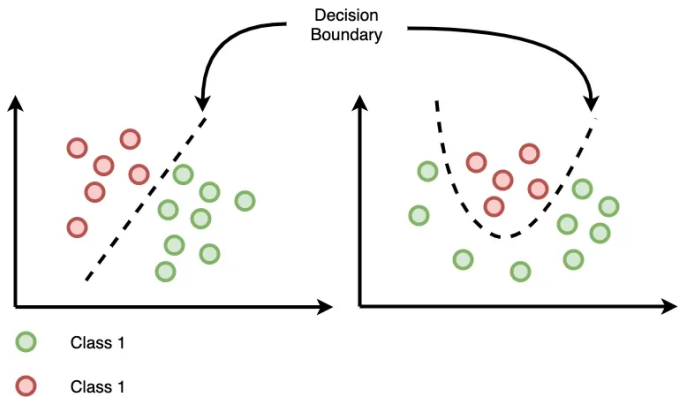
Parametreler; $z = \theta^T x$

Lojistik Fonksiyon; $g(z) = \frac{1}{1+e^{-z}}$

Yukarıda görüldüğü gibi $h(x) = g(z)$, 0 ve 1 aralığındaki çıktı değerlerini veya olasılıkları hesaplamak için kullanılan hipotez fonksiyonu veya lojistik fonksiyonudur. Lojistik fonksiyonumuzun davranış şekli, girdisi olan z sıfıra eşit veya sıfırdan büyük olduğunda, çıktısı $g(z)$ 0,5'e eşit veya 0,5'ten büyük olmasıdır. İkili sınıflandırmada ise çıktımızın ya 1 ya da 0 olması gerekir, bu nedenle doğrusal regresyondan farklı olarak hipotezimizin de 0 ve 1 aralığında olması gerekmektedir, hipotez için de lojistik fonksiyon kullanmamızdaki amaç budur. Genellikle ulaştığımız 0 ve 1 arasındaki değer olasılık olarak yorumlanmaktadır. Örneğin, gerçek veya spam e-posta sınıflandırma problemi için oluşturduğumuz modelde $h(x)$ hipotezinin çıktı değerinin $h(x)$ = 0,83 olduğu sonucuna ulaşıldığında burada 0 gerçek e-postayı ve 1 ise spamı temsil etmektedir. Ulaştığımız sonuç basitçe, e-postanın spam olma olasılığının yüzde seksen üç olduğu anlamına gelmektedir. Lojistik regresyon doğrusal karar sınırlarının belirlenmesinde kullanılabilen bir algoritmadır. Karar sınırı veri kümesini çeşitli

sınıflarına ayıran bir sınırdır. İkili sınıflandırma probleminde karar sınırı, veri kümesindeki pozitif örnekleri negatif örneklerden ayıran bir çizgi olarak karşımıza çıkmaktadır.

Şekil 1.5. Karar sınırları örnekleri



Kaynak: (Rostamizadeh & Talwalkar, 2012)

Lojistik regresyon kullanılarak doğrusal karar sınırları belirlenebilir ancak doğrusal olmayan karar sınırları oluşturarak daha doğru sonuçlar elde edilebilecek modeller oluşturulması da mümkündür.

1.2.2. Yapay sinir ağları

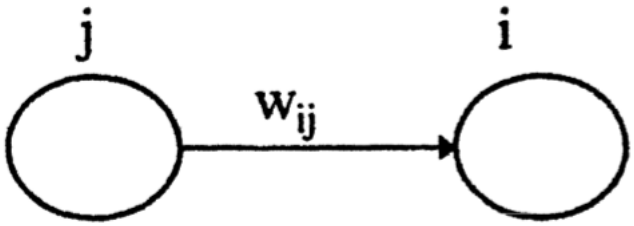
Makine öğrenmesi kapsamında yapay sinir ağı olarak tanımlanan konsept temel elemanlar arasında gerçekleştirilen basit işlemler bütünü olarak tanımlanabilmektedir. Her ne kadar isimlendirme doğrudan canlı sinir sistemini temel alsa da çalışma şekilleri benzerlik göstermemektedir. Canlı sinir sisteminde olduğu gibi bu sistemdeki temel yapı taşı nöron olarak adlandırılmaktadır. Canlı nöronlar ve makine öğrenmesi kapsamında kullanılan nöronlar karmaşıklık ve çalışma şekilleri olarak kıyaslanmamalıdır(Taylor, 1993).

Nöronların her biri belirlenen aralıkta bir değer tutan, bu değer bir vektör, sayı olabilir bir birim olarak düşünülebilir. Genellikle bu değer 0 ve 1 dahil bu sayılar arasında olacak şekilde normalleştirilmektedir. Bu değer nöronun bağlı olduğu diğer nöronlara

gönderilecek sinyaldir. Sinyaller nöronun bağlı bulunduğu ağdaki komşu nöronlarına gönderilir. Her nöron komşusundan gelen sinyali almaktadır. İki nöron arasındaki bağlantıda "bağlantı ağırlıkları" bulunmaktadır, gönderilen sinyal bu ağırlıklara göre değiştirilmektedir. i ve j isimli birbirine bağlı iki nöron bulunuyorsa bu nöronlar arasındaki bağlantının ağırlığı w_{ij} olarak adlandırılmaktadır. Dolayısıyla, j nöronun i nöronuna göndereceği 1 değerinde bir sinyal i nöronunu w_{ij} değerinde aktive edecektir. i nöronun toplam aktive değeri şu şekilde gösterilebilir;

$$A_i = \sum_i w_{ij} u_j$$

Şekil 1.6. İki Nöronlu Sistem



Kaynak: (Taylor, 1993)

Burada u_j j nöronunun aktivasyon değerini göstermektedir. Bu değer deterministik olabilir yani sadece 0 ya da 1 değeri alabilir, olasılıksal bir şekilde değerlendirilebilir bu durumda A_i değeri sigmoid bir fonksiyonda değerlendirilerek olasılığa dönüştürülmektedir ve sonucunda nöronun aktive olma olasılığı belirlenmektedir. Son olarak ise aktivasyon değeri gerçek bir sayı olarak da değerlendirilebilmektedir(Taylor, 1993).

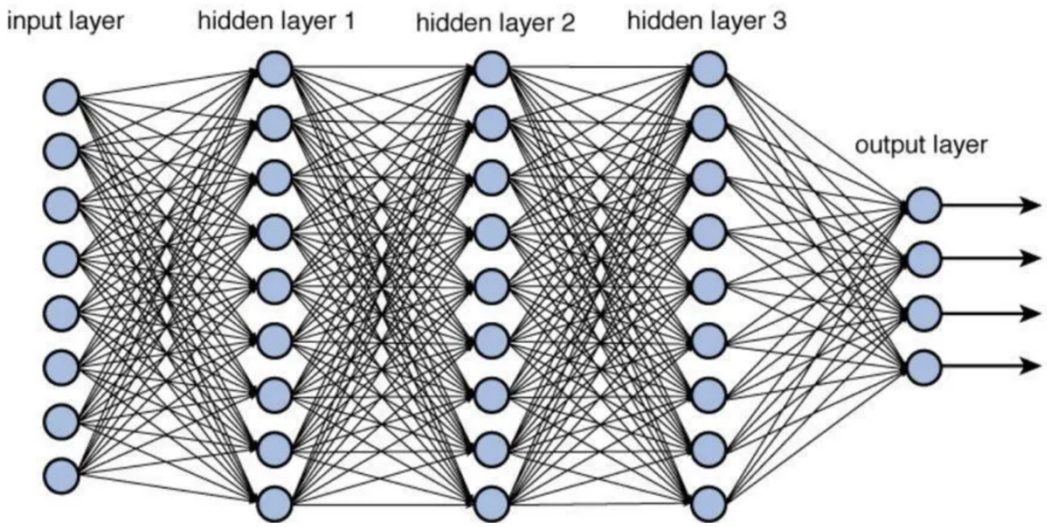
Tablo 1.3. Nöron Çeşidi Çıktı Tablosu

Nöron çeşidi	**Çıktı**
Deterministtik	1 eğer $A_{\mathrm{i}} > 0$, 0 eğer $A_{\mathrm{i}} < 0$
Olasılıksal	1 olma olasılığı $f(A_i)$
Gerçek sayı	$f(A_i)$

Kaynak:(Taylor, 1993)

Yapay sinir ağlarında veriyi girdi olarak verdiğimiz nöronlara giriş nöronları, verinin işlendiği nöronlara saklı nöronlar, verinin işlenip son halini aldığı nöronlara ise çıkış nöronları denmektedir(Taylor, 1993). Girdi ve çıktının aynı boyutta olması gerekmemektedir. Örneğin, sınıflandırma sorularında çok boyutlu bir girdi alıp sadece pozitif sonuçlarda aktive olan tek bir çıktı kullanabilmektedir. Nöronların sıralandığı her bir gruba katman denmektedir.

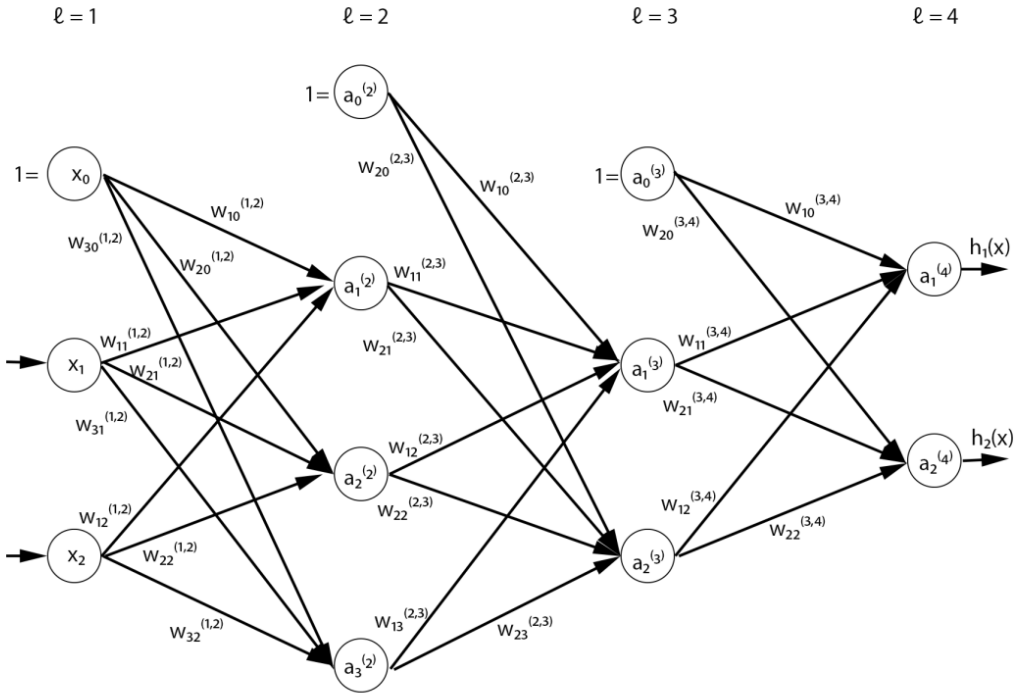
Şekil 1.7. Çok Katmanlı Sinir Ağı



Kaynak:(Parmar, 2024)

Dört katmanlı bir yapay sinir ağımız olduğunu varsayalım birinci katman $\ell = 1$ girdi katmanımız, ikinci ve üçüncü katmanımız olan $\ell = 2$ ve $\ell = 3$ saklı katman, son katmanımız olan $\ell = 4$ ise çıktı katmanımız olmaktadır. Bu örnekte girdi katmanının saklı katman ile saklı katmanın bir diğer saklı katman ile ve son olarak saklı katmanın çıktı katmanı ile bağlantısı bulunmaktadır. Bu tüm bağlantı çeşitlerine sahip en sade yapay sinir ağı örneğidir.

Şekil 1.8. Yanlı/Sapma Nöronları Olan Çok Katmanlı Bir Sinir Ağı



Kaynak:(Bougourzi & Bougourzi, 2020)

$\ell = 1,2, \dots, N$ katmanı tanımlar N toplam katman sayımızdır. Yukarıdaki örnekte $N = 4$. Çıktı katmanı hariç her katmanda bulunan ve sürekli 1 değeri alan nöronlara yanlı/sapma nöronları denmektedir. Bu nöronlar modele esneklik kazandırmak ve verilere uyumu arttırmak için eklenmektedir, genelde, tüm girdiler 0'a eşit olduğunda ağın verileri işlemesine olanak tanır ve çoğu durumda ağ içinde takılan değerlerin sapmasını azaltır(Robert Keim, 2020). n_ℓ bir katmanda yanlı sapma nöronları hariç kaç nöron olduğunu belirtmektedir. Örnekte $n_1 = 2,\ n_2 = 3,\ n_3 = 2,\ n_4 = 2$

olmaktadır, nöronlar arasındaki ağırlıklar $w_{ij}^{(\ell,\ell+1)}$ ile ifade edilmektedir. Bu ifade ℓ katmanının j nöronu ve $\ell+1$ katmanının i nöronu arasındaki bağlantının ağırlığını temsil etmektedir. En sağda bulunan x_1 ve x_2 yapay sinir ağımızı beslediğimiz girdi verileridir. En solda bulunan $h_1(x)$ ve $h_2(x)$ ise çıktı katmanının verdiği sonuçları temsil etmektedir. a_i^{ℓ} ise ℓ katmanındaki i nöronunun aktivitesini temsil etmektedir(Bougourzi & Bougourzi, 2020).

Bir yapay sinir ağının eğitilmesi demek sonuçları önceden belirlenmiş girdi ve çıktılar kullanılarak modelde bulunan parametrelerin ağırlıkların en doğru cevabı sağladığı durumda optimum hale getirilmesi demektir. Eğitim işlemi döngüsel olup bilgisayar kaynağı açısından pahalı bir işlemdir. Eğitim makine öğrenmesi kısmında bahsedilen maliyet fonksiyonlarının ağırlıklara göre kısmi türevleri alınarak gerçekleştirilmekte ve bilgisayar için işlem gücü maliyeti fazla olan lineer cebir operasyonlarının yapılmasını içermektedir. Eğitim sırasında sıklıkla kullanılan bu yöntemin adı geri yayılım algoritmasıdır.

Yapay sinir ağı mimarisinde kaç girdi, kaç çıktı, kaç nöron ve kaç katmana ihtiyaç duyacağınız çözmek istediğiniz soruna ve sorunu çözmek için kullandığınız veriye göre değişmektedir. Örnek olarak 1000 x 1000 çözünürlükte hayvan resimlerinden kedi ve köpekleri ayırt edecek bir yapay sinir ağı geliştirmeyi hedefliyorsak her pikselin renk değerini alacak şekilde bir milyon girdi nöronu, kedi, köpek ve diğer olacak şekilde 3 çıktı nöronuna ihtiyaç duyulabilir. Araya yerleştireceğimiz her saklı katman modelimizin performansını arttıracaktır, fakat bununla birlikte eğitim için gereken hesaplama miktarı ve eğitim süresi de orantılı bir biçimde artmaktadır.

1.2.2.1 Derin öğrenme

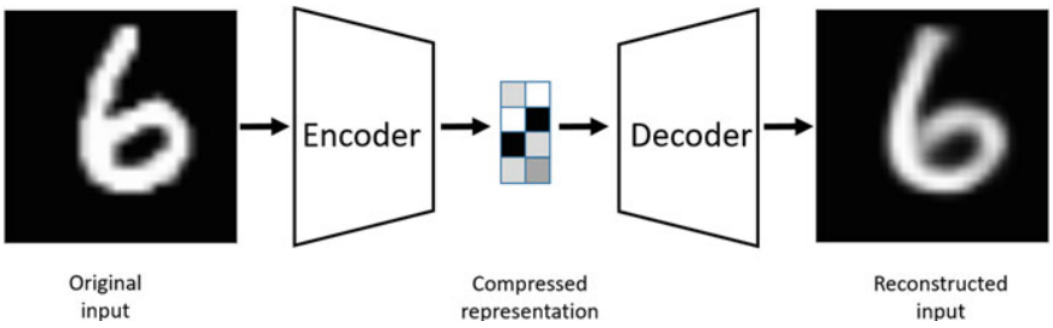
Derin öğrenme makine öğrenmesinin en popüler ve günümüzde en çok bahsi geçen alt dallarından biridir. Yapay sinir ağlarının geliştirilmiş versiyonu olarak karşımıza çıkmaktadır. Derin yapay sinir ağları gibi bu derin öğrenme yöntemleri, çok büyük veri kümelerindeki örüntüleri ve yapıları keşfetmek için birden fazla işlem katmanı

kullanır. Her katman, sonraki katmanların üzerine inşa ettiği verilerden bir kavram öğrenir. Basit bir örnek vermek gerekirse, şekilleri yorumlamakla görevli derin bir sinir ağı, ilk katmanda basit kenarları tanımayı öğrenecek ve ardından sonraki katmanlarda bu kenarlardan oluşan daha karmaşık şekillerin tanınmasını ekleyecektir. Derin öğrenme için kaç katman gerektiğine dair kesin ve hızlı bir kural yoktur, ancak çoğu uzman ikiden fazla katman gerektiğini belirtmektedir(Rusk, 2015).

Kısacası, derin öğrenme, çok katmanlı yapay sinir ağıdır ve derin sinir ağı olarak da adlandırılabilir. Veri girişine yakın alt katmanlar basit özellikleri öğrenirken, üst katmanlar alt katman özelliklerinden türetilen daha karmaşık özellikleri öğrenir. Mimari, hiyerarşik ve güçlü bir özellik temsili oluşturur. Bu, derin öğrenmenin hem büyük miktarda veriden hem de farklı kaynaklardan toplanan verilerden yararlı bilgileri analiz etmek ve çıkarmak için uygun olduğu anlamına gelir(L. Zhang vd., 2018).

Makine öğrenmesinin diğer alanlarında olduğu gibi derin öğrenme alanında da pek çok model kullanılmaktadır. Yaygın olarak kullanılan bazı modeller arasında Otokodlayıcı (*Autoencoder* - AE), Derin İnanç Ağları (*Deep Belief Network* - DBN), Evrişimli Sinir Ağları (*Convolutional Neural Network* - CNN) ve Yinelemeli Sinir Ağı (*Recurrent Neural Network* - RNN) bulunmaktadır.

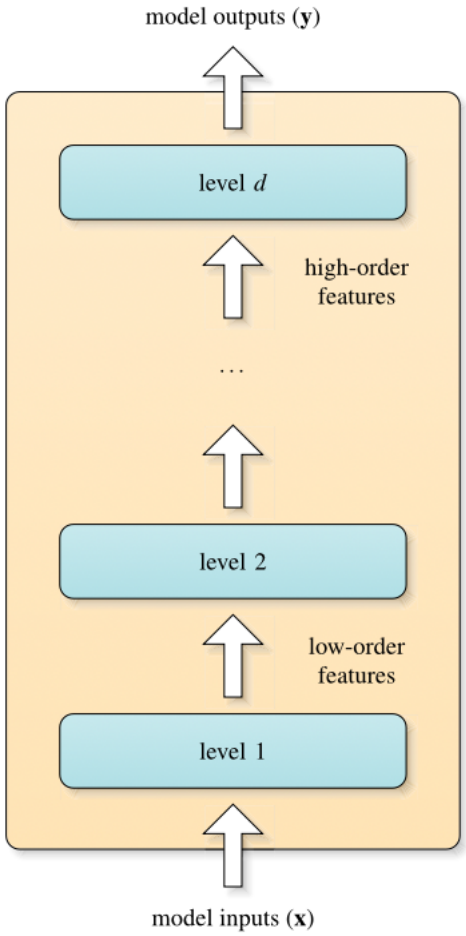
Şekil 1.9. AE Mimarisi



Kaynak: (Rokach vd., 2023)

AE girdisini yeniden yapılandırmak için eğitilen bir sinir ağı olarak tanıtılmıştır. Temel amaçları, kümeleme gibi çeşitli görevler ve çıkarımlar için kullanılmak üzere denetimsiz bir şekilde öğrenmektir(Rokach vd., 2023). AE, etiketlenmemiş veriler üzerinden eğitilen bir yapay sinir ağı türüdür. İki işlevi öğrenir: giriş verilerini dönüştüren bir *encoder* ve giriş verilerini kodlanmış temsilden yeniden oluşturan bir *decoder*(Kramer, 1991). AE genellikle, *encode* sırasında çok boyutlu girdilerin daha az boyutlarda temsil edilmesi yöntemi ile çalışmaktadır.

Şekil 1.10. DBN mimarisi

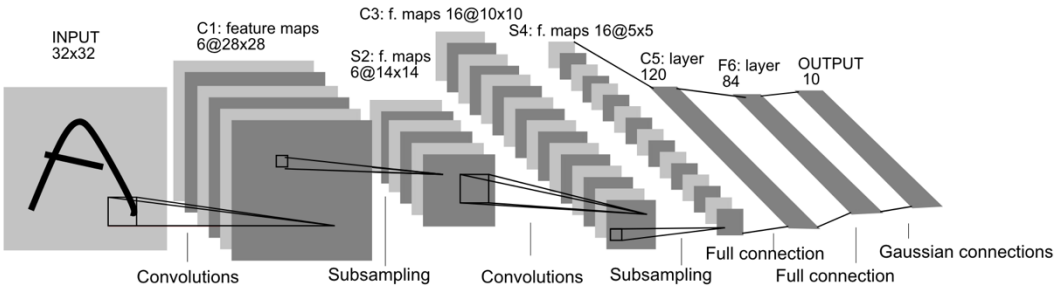


Kaynak: (Lopes & Ribeiro, 2015)

DBN çok katmanlı fakat katmanları oluşturan nöronlar arasında bağlantı olmadan doğrudan katmanların çıktılarının aktarıldığı bir yapay sinir ağı modelidir. Her katman için bir önceki katmandaki birimlerin istatistiksel bağımlılıkları, bağımsız olarak

eğitilen bir yapay sinir ağıdır. Bir DBN, eğitim verilerinin olasılığını en üst düzeye çıkarmayı amaçladığından, eğitim süreci DBN girdilerini alan alt düzey yapay sinir ağları ile başlar ve hiyerarşide kademeli olarak yukarı doğru hareket eder, sonunda DBN çıktılarını içeren en üst katmandaki yapay sinir ağı eğitilir(Lopes & Ribeiro, 2015).

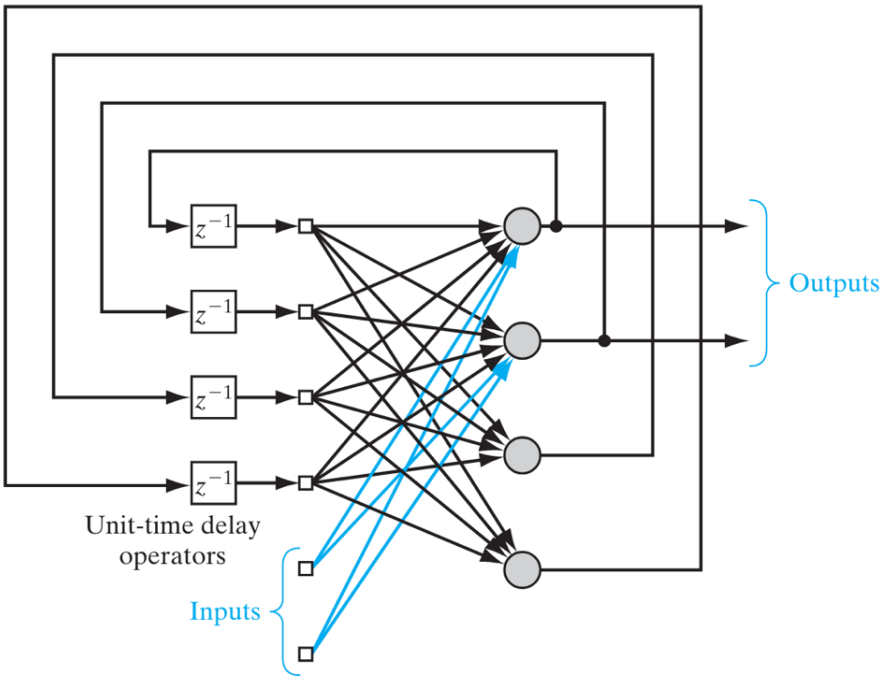
Şekil 1.11. CNN Mimarisi



Kaynak:(Lecun vd., 1999)

CNN, canlıların görsel sisteminden esinlenen özel bir tür çok katmanlı sinir ağı veya derin öğrenme mimarisidir. CNN, bilgisayarla görme ve doğal dil işleme gibi alanlarda sıklıkla kullanılmaktadır(Balas vd., 2020). Geleneksel CNN, bir veya birden fazla evrişim (*convolution*) ve havuzlama (*pooling*) katmanı bloğunun ardından bir veya birden fazla tam bağlı (*fully connected* - *FC*) katman ve bir çıkış katmanından oluşmaktadır. CNN'ler ilk olarak LeCun ve diğerleri tarafından önerilmiştir(Lecun vd., 1999), LeNet adı verilen çok temel ve popüler bir CNN ağı aşağıda yer almaktadır.

Şekil 1.12. RNN Mimarisi



Kaynak: (Haykin, 2009)

RNN ya da tekrarlayan sinir ağı kendi çıktısı ile tekrar kendini besleyerek diğer sinir ağlarından ayrılmaktadır. Yukarıda saklı katmanları da olan bir tekrarlayan sinir ağı örneği yer almaktadır. Geri bildirim sistemi sinir ağının öğrenme kabiliyeti ve performansı üstünde etkiye sahiptir. Ayrıca şekilde z^{-1} ile gösterilen birim zamanlı gecikme elemanları ile lineer olmayan bir sinir ağı elde edilebilmektedir(Haykin, 2009).

Verilen girdileri tekrar işleme özellikleri yazı ve konuşma algılama gibi uygulamalarda RNN yönteminin sıkça tercih edilmesini sağlamaktadır(Sak vd., 2014).

2. ULUSLARARASI KURULUŞLARIN YAPAY ZEKA ÇALIŞMALARI

Yapay zeka, son yıllarda hızla gelişen ve geniş bir etki alanına sahip olan bir teknoloji olarak öne çıkmaktadır. Bu doğrultuda uluslararası kuruluşlar, yapay zekanın etik, güvenli ve etkili bir şekilde kullanılmasını sağlamak amacıyla çeşitli çalışmalar yürütmektedir. Bu kuruluşlar, yapay zeka politikaları geliştirmekten, standardizasyon çalışmalarına kadar geniş bir yelpazede faaliyet göstermekte ve yapay zekanın toplumsal, ekonomik ve güvenlik alanlarındaki etkilerini incelemektedir. Avrupa Birliği (*European Union*-EU/AB), ITU, Ekonomik İşbirliği ve Kalkınma Örgütü (*Organisation for Economic Co-operation and Development*-OECD), Avrupa Elektronik Haberleşme Düzenleyicileri Kurumu (*Body of European Regulators for Electronic Communications*-BEREC), Birleşmiş Milletler (*United Nations*-UN) ve NATO gibi kuruluşlar yapay zeka konusunda çalışmalar gerçekleştirmektedir.

2.1. Avrupa Birliği

Avrupa Komisyonu, yapay zekanın güvenilir, etik ve rekabetçi bir şekilde geliştirilmesi için 2020 yılında bir Beyaz Kitap yayımlamıştır. Bu doküman, AB'nin yapay zeka stratejisinin temel taşlarını oluşturmaktadır ve etik kurallar, güvenlik önlemleri ve inovasyon teşvikleri gibi konulara değinmektedir(European Commission, 2020). Bu rapor ile Avrupa Komisyonu, yapay zekanın gelişimini ve kullanımını teşvik ederken, aynı zamanda bu yeni teknolojinin kullanımına ilişkin riskleri ele almayı amaçlamaktadır. Düzenleyici ve yatırım odaklı bir yaklaşım desteklenmiştir. Beyaz Kitap, bu hedeflere nasıl ulaşılacağına dair politika seçenekleri sunmaktadır ve gelecekteki karar verme süreçlerine katkıda bulunmak üzere tüm paydaşların görüş bildirmesini talep etmektedir(European Commission, 2020).

AB'nin gerçekleştirdiği bir diğer yapay zeka çalışması ise Avrupa Yapay Zeka Platformu olan AI4EU'dur. Yapay zeka teknolojilerini paylaşmak, geliştirmek ve işbirliklerini artırmak amacıyla kurulmuştur. Platform, araştırma projeleri, veri kaynakları ve eğitim materyalleri gibi birçok kaynağı bir araya getirmektedir(European Union, 2024). AI4EU projesi, 2019'da Avrupa

Komisyonu'nun H2020[1] programı desteğiyle kurulmuştur. Proje, AB ülkelerindeki bilim insanları, girişimciler, KOBİ'ler, sanayiler, finansman kuruluşları ve vatandaşlar arasında işbirliğini artırarak geniş bir ekosistem oluşturmayı hedeflemektedir.

En son olarak ise AB Komisyonu tarafından 21 Nisan 2021 tarihinde sunulan ve yapay zekâ sistemlerinin piyasaya arzı, hizmete sunulması ve bazı uygulamaların yasaklanmasına dair kuralları belirleyen "Yapay Zekâ Hakkında Uyumlaştırılmış Kurallar Getiren ve Bazı Birlik Yasama Tasarruflarını Değiştiren Yasa Önerisi", AB Konseyi'nde resmi olarak onaylanmıştır. AB bu düzenleme ile yapay zeka teknolojilerinin Avrupa değerlerine, demokratik toplum yapısına ve temel haklara zarar vermeden ekonomiye ve topluma yapacağı katkıları en üst seviyeye taşımayı hedeflemektedir. Bu çerçevede risk temelli bir yaklaşımla hazırlanan Yasa, topluma zarar verme potansiyeli yüksek olan insan davranışlarını yönlendirme, insanların zafiyetlerini kullanma veya biyometrik sınıflandırma gibi işlevlere sahip yapay zeka sistemlerinin kullanımını yasaklamaktadır(Eprs & Rapporteurs, 2024).

Bu düzenleme yapay zeka sistemlerinin AB pazarında kullanımına ve tedarikine yönelik ilk bağlayıcı metindir. Yasa, risk bazlı bir yaklaşım benimseyerek yapay zeka sistemlerini farklı risk seviyelerine göre sınıflandırmakta ve çeşitli yükümlülükler getirmektedir. "Kabul edilemez" risk taşıyan yapay zeka sistemleri yasaklanırken, "yüksek risk" içeren sistemler belirli gerekliliklere tabi tutulmaktadır. Az riskli sistemler bilgi ve şeffaflık gereksinimlerine tabi olmaktadır. Minimal riskli sistemler ise ek yükümlülüklerden muaf tutulmaktadır. Ayrıca, genel amaçlı yapay zeka modelleri için özel kurallar ve yüksek etki kapasitesine sahip olanlar için daha sıkı gereklilikler belirlenmiştir(European Parliament, 2024).

[1]H2020, tam adıyla Horizon 2020, Avrupa Birliği'nin 2014-2020 yılları arasında yürüttüğü araştırma ve inovasyon programıdır.

2.2. Ekonomik İşbirliği ve Kalkınma Örgütü

OECD, yapay zeka kullanımına yönelik ilk uluslararası politikayı geliştirmiştir. Bu prensipler, insan haklarına saygı, adalet, şeffaflık ve hesap verebilirlik gibi temel değerleri içermektedir. OECD Yapay Zeka İlkeleri (*AI Principles*) ilk olarak 2019'da kabul edilmiş ve Mayıs 2024'te güncellenmiştir(OECD, 2024). Üye ülkelerin katkısıyla, yeni teknolojik ve politik gelişmeler göz önünde bulundurularak, bu ilkeler güncellemiş ve amaca uygun kalması sağlanmıştır.

İlkeler, yapay zeka aktörlerine güvenilir yapay zeka geliştirme çabalarında rehberlik etme ve politika yapıcılar için etkili yapay zeka politikaları oluşturma konusunda öneriler sunmayı amaçlamaktadır. Ülkeler, politikalar oluşturmak ve yapay zeka risk çerçeveleri oluşturmak için OECD Yapay Zeka İlkeleri ve ilgili araçları kullanmaya ve bu sayede uluslararası uyumluluk ve işbirliği sağlamaya davet edilmektedir(OECD, 2024).

OECD ayrıca, üye ülkelerin yapay zeka politikalarını takip etmek ve analiz etmek için bir gözlem merkezi kurmuştur. Bu merkez, yapay zekanın ekonomik ve sosyal etkilerini incelemekte ve üye ülkelere politika önerilerinde bulunmaktadır(OECD, 2024).

2.3. Uluslararası Telekomünikasyon Birliği

ITU, telekomünikasyon ve bilgi teknolojileri alanında yapay zeka ve makine öğrenimi hakkında çalışmalar gerçekleştirmektedir. Telekomünikasyon Standardizasyon Sektörü (*Telecommunication Standardization Sector*-ITU-T) bünyesinde çalışmalar yürüten 16. Çalışma Grubu ve Radyo-iletişim Sektörü (Radiocommunication Sector-ITU-R) bünyesinde çalışmalar yürüten 6. Çalışma Grubu yapay zeka konusunda da çalışmalar yürütmektedir. Bu çalışma grupları, görsel/işitsel uygulamalar, multimedya bilgi sistemleri ve telekomünikasyon altyapıları üzerine odaklanmakta ve kendi alanlarında yapay zeka ve makine öğrenmesi uygulamaları üzerine çalışmalar gerçekleştirmektedirler(ITU, 2024a).

Ayrıca ITU tarafından *AI for Good* isimli bir zirve düzenlenmektedir. Bu zirve, yapay zekanın toplumsal faydalarına yönelik küresel bir diyalog ortamı oluşturmayı amaçlamaktadır. Zirvede, yapay zekanın sağlık, eğitim ve çevre gibi çeşitli alanlarda nasıl kullanılabileceği tartışılmaktadır(ITU, 2024b).

2.4. Avrupa Elektronik Haberleşme Düzenleyicileri Kurumu

BEREC 2024 yılında yapay zeka ve sanal dünyalar konularında duruşunu belirleyen bir belge yayınlamıştır. Bu belgede, yapay zeka ve sanal dünyaların dijital ekonomideki artan önemine ve bu teknolojilerin başarılı bir şekilde geliştirilip uygulanabilmesi için yüksek kaliteli elektronik iletişim ağları ve hizmetlerine ihtiyaç duyduğuna vurgu yapılmıştır. Özellikle yüksek kapasiteye sahip ağlar, bulut hizmetleri ve uç bilişimin bu teknolojiler için kritik öneme sahip olduğu vurgulanmıştır(BEREC, 2024).

BEREC raporunda, yapay zeka konusunda rekabet dinamikleri, çevresel ayak izi ve sürdürülebilirlik konuları gibi bazı temel sorunların ele alınması gerektiğini belirtmiştir. Yayınlanan raporda BEREC, yapay zekanın hesaplama gücü, veri, finansal kaynaklar ve teknik uzmanlık şeklinde dört temel girdiyle mümkün olduğunu belirtmektedir. Bu girdilere ayrıcalıklı veya münhasır erişimin, önemli bir rekabet avantajı sağlayabileceği ve sektöre giriş konusunda engeller oluşturabileceği belirtilmiştir. Son olarak, raporda yapay zekanın siber güvenlik üzerindeki etkileri de ele alınmaktadır. Yapay zekanın, güvenlik tehditlerini artırabilirken, aynı zamanda daha güçlü güvenlik önlemleri geliştirilmesine de yardımcı olabileceği belirtilmiştir.

2.5. Birleşmiş Milletler

BM'nin eğitim, bilim ve kültür örgütü (*United Nations Educational, Scientific and Cultural Organization*-UNESCO), yapay zekanın etik kullanımına yönelik rehberler ve politikalar geliştirmektedir. UNESCO, yapay zekanın eğitimde kullanımı ve dijital becerilerin geliştirilmesi üzerine de çalışmalar yapmaktadır. Örneğin, yapay zekanın

eğitimde nasıl etkili bir şekilde kullanılabileceğine dair raporlar ve kılavuzlar yayınlamaktadır(United Nations, 2024).

BM'nin ayrıca *UN Global Pulse* isimli bir girişimi bulunmaktadır. Bu girişimi, büyük veri ve yapay zekayı kullanarak küresel insani yardım ve kalkınma çalışmalarını desteklemektedir. *UN Global Pulse*, veri analitiği ve yapay zeka teknolojilerini kullanarak, kriz müdahalesi ve sürdürülebilir kalkınma hedeflerine ulaşmayı amaçlamaktadır(United Nations, 2024a).

Ayrıca Birleşmiş Milletler içinde kurulmuş bir Yapay Zeka Yüksek Düzeyli Danışma Kurulu bulunmaktadır. Kurulun amacı yapay zeka ile ilgili hizmetler, algoritmalar, hesaplama kapasitesi ve uzmanlığın uluslararası alanda daha yaygın hale gelmesiyle birlikte gelen faydanın risklerini ve belirsizliklerini ele almak ve yapay zeka yönetiminin küresel olarak koordine edilmesine yardımcı olmaktır. Dünya genelinde ilgili disiplinlerdeki otuzdan fazla uzmanın bir araya gelmesiyle oluşturulan Kurul, yapay zekanın ortak iyilik için nasıl yönetilebileceğine dair çeşitli perspektifler ve seçenekler sunmayı amaçlamaktadır. İnsan hakları ve sürdürülebilir kalkınma hedefleri ile uyumlu uluslararası birlikte çalışabilir bir yapay zeka yönetimi sağlanması hedeflenmektedir (United Nations, 2024b).

2.6. Kuzey Atlantik Antlaşması Örgütü

Yapay zeka, NATO Müttefiklerinin savunma ve güvenlik açısından önem verdikleri yedi teknolojik alandan biridir. Bunlar arasında kuantum destekli teknolojiler, veri ve hesaplama, otonomi, biyoteknoloji ve insan geliştirmeleri, hipersonik teknolojiler ve uzay bulunmaktadır. 21 Ekim 2021 tarihinde NATO Savunma Bakanları NATO'nun ilk Yapay zeka stratejisini kabul etmiştir. Strateji, yapay zekanın savunma ve güvenlik alanında güvenli ve etik bir şekilde nasıl uygulanabileceği üzerine durmaktadır. Bu bağlamda, uluslararası hukuk ve NATO'nun değerlerine uygun olarak yapay zeka teknolojilerinin sorumlu kullanımına dair standartlar belirlemektedir. Ayrıca, hasımların yapay zeka kullanımıyla ortaya çıkan tehditleri ele almakta ve yapay zeka

konusunda çalışmalar yürüten müttefikler arasında güvenilir işbirliği kurmanın yollarını belirlemektedir(NATO, 2024).

Stratejinin dört temel amacı bulunmaktadır;

1. NATO ve müttefikler için örnek teşkil etmek ve müttefik savunma ve güvenlik amaçları için yapay zekanın sorumlu bir şekilde geliştirilmesini ve kullanılmasını teşvik etmek;
2. Yapay zeka konusunda yetenek geliştirme ve teslimatını hızlandırmak ve ana akıma taşımak, ittifak içinde birlikte çalışabilirliği artırmak;
3. Müttefiklerin yapay zeka teknolojilerini ve yenilik yapma kabiliyetini korumak ve izlemek;
4. Devlet ve devlet dışı aktörlerin kötü niyetli YZ kullanımından kaynaklanan tehditleri belirlemek ve koruma sağlamak.

3. YAPAY ZEKANIN SİBER GÜVENLİK ALANINDA KULLANIMLARI VE ÜLKE İNCELEMELERİ

Yapay zeka insan beynini taklit etmek ve gerçek dünya problemlerini bir insan yaklaşımıyla makinelere çözdürmek ya da nasıl çözdürüleceğini araştırmak için ortaya atılmış bir kavramdır. Bilgi teknolojisi ve fizyolojik zekanın bir kombinasyonu olarak özetlenebilmektedir. Bu sistemler büyük miktarda verinin akıllıca kullanılmasını ve işlenmesini mümkün kılarak işlevsel araçların oluşturulmasını sağlamaktadır. Yapay zeka, savunma, sağlık ve uzay araştırmaları gibi çeşitli alanlarda kullanılmaktadır.

Buna benzer olarak, günümüzde giderek daha yetenekli hale gelen saldırganlara karşı bilgi güvenliğini sağlamak için siber güvenlik alanında da yapay zekadan yararlanılmaktadır. Günümüzde Yapay zeka, saldırıları ve bilgi sistemleri ihlallerini tespit etmek ve tepki vermek için oluşturulan karmaşık süreçlerin otomatikleştirilmesine yardımcı olmaktadır. Bu tür uygulamalar, yapay zekadan yararlanarak gelişmekte ve daha gelişmiş ve kapsamlı hale gelmektedir. Makine öğrenmesi, yapay zekanın temel bir alanıdır. Bilgisayarlara deneyim yoluyla öğrenmeyi ve uyum sağlamayı öğretmek için bir araç sağlayan teknolojileri ifade eder. Bu teknoloji, akıl yürütme (neden ve sonuç) yerine deneyim ve kalıplardan öğrenme yöntemlerinden faydalanmaktadır.

Siber güvenlik, risklerin yönetilmesi, güvenlik açıklarının giderilmesi ve sistem direncinin artırılması olarak tanımlanabilir. Siber güvenlikte, ağ davranış anormallikleri ve kötü amaçlı yazılımların tespit edilmesiyle ilgili teknikler de yer almaktadır. Siber güvenlik, siber saldırılara ve bunların sonuçlarına karşı savunmada alınan bir dizi eylem ve gerekli karşı önlemlerin uygulanması olarak tanımlanabilmektedir (Prasad & Rohokale, 2015).

Siber saldırılar, tek bir eylemden veya birbiriyle bağlantılı ayrı adımların birleşiminden oluşabilen birçok biçimde gerçekleştirilebilmektedir. Bu tür eylemler, karmaşık mühendislik yöntemleri ile sistemlerin istismar edilmesi ya da gizli bilgilere

erişim sağlamak için sosyal mühendisliğin basit bir şekilde kullanılması ile gerçekleştirilebilmektedir.

Yapay zeka ve makine öğrenmesi, siber saldırıları tespit etmek, bunlara karşı savunma yapmak ve incelemek için kullanılabilmektedir. İnsanlar tarafından oluşturulan ve analitik olarak adlandırılan çözümler, siber güvenlik uzmanlarının belirlediği kurallara dayanır ve bu kurallara uymayan saldırılar gözden kaçabilmektedir. Bunun yanı sıra makine öğrenmesi tabanlı yaklaşımlar ise, yanlış pozitif sonuçlar oluşturabilmekte ve dolayısıyla vakaları araştırmak için tekrardan insan müdahalesi gerektirebilmektedir. Sektörün giderek büyümesi ve sistemlerin karmaşıklaşmasına paralel olarak oluşan büyük miktarda verinin siber güvenliğini yönetmek için etkili yapay zeka çözümlerine ihtiyaç duyulmaktadır(Sipola vd., 2023).

3.1. Ülkelere Göre Yapay Zeka Siber Güvenlik Uygulamaları

3.1.1. Amerika Birleşik Devletleri

Makine öğrenmesi, ABD'de siber güvenlik önlemlerinin geliştirilmesinde çok önemli bir araç haline gelmiştir. Hem kamu hem de özel sektörde çok sayıda uygulama geliştirilmiş olup dünya çapında kullanılmaktadır.

Siber Güvenlik ve Altyapı Güvenliği Ajansı (*Cybersecurity and Infrastructure Security Agency* - CISA), Amerika Birleşik Devletleri İç Güvenlik Bakanlığı'nın (*United States Department of Homeland Security* - DHS), devletin tüm kademelerinde siber güvenlik ve altyapı korumasından, eyaletlerin siber güvenlik programlarının koordine edilmesinden ve hükümetin her türlü bilgisayar korsanlığı faaliyetine karşı korunması ve siber güvenlik çözümleri geliştirmesinden sorumlu ajansıdır(CISA, 2024a). Kamu tarafından geliştirilen yapay zeka destekli siber savunma projelerinin bir çoğu CISA tarafından geliştirilmektedir. Şöyle ki;

- **Otomatik Gösterge Paylaşımı (*Automated Indicator Sharing*-AIS) Puanlama ve Geri Bildirim (*Scoring & Feedback*-AS&F)**

  Bir CISA uygulaması olan Otomatik Gösterge Paylaşımı (AIS), siber olaylara karşı korunma ve bu olayların yaygınlığını azaltmak için makine tarafından

okunabilir siber tehdit göstergelerinin ve savunma önlemlerinin gerçek zamanlı paylaşımını sağlamaktadır. AIS Çerçevesi üzerine inşa edilen AS&F, AIS aracılığıyla paylaşılan tehdit bilgilerinin, bilginin görüşü gönderen taraf için mevcut olan diğer kaynaklarla bilginin doğrulanıp doğrulanamayacağına dair bir değerlendirme sağlayan bir görüş değeri ve gönderenin AIS'ye gönderdiği bilgilerin doğruluğuna olan güvenini belirten bir güven puanı ile zenginleştiren bir algoritmadır. CISA tarafından kullanıldığında, AS&F yapay zeka ve makine öğrenmesini kullanarak gelen istihbaratın ilk analizlerini gerçekleştirmektedir(CISA, 2024b).

- **AIS Otomatik Kişisel Olarak Tanımlanabilir Bilgi (*Personally Identifiable Information*-PII) Tespiti**

  Otomatik PII Tespiti, Otomatik Gösterge Paylaşımı gönderimlerindeki olası PII verisini tespit etmek için doğal dil işleme sistemlerinden yararlanan bir sistemdir. Gönderilerde olası PII tespit edilirse, söz konusu gönderiler insan incelemesi için ayrılmaktadır. İnsan incelemesi kapsamında, analistler PII verisinin doğruluğunu onaylayabilir/reddedebilir ve gerekirse bilgileri çıkartabilir. Sistem, analistlerden ve uzmanlarından gelen geri bildirimlerden öğrenmektedir(CISA, 2024b).

- **Gelişmiş Analitik Destekli Adli Soruşturma (*Advanced Analytic Enabled Forensic Investigation*)**

  CISA, Federal Sivil Yürütme Organı (*Federal Civilian Executive Branch* - FCEB) departmanları, ajansları ve kritik altyapı ortaklarında gerçekleşen siber olayları analiz etmek için adli bilişim uzmanları görevlendirmektedir. Bu uzmanlar, anomalileri ve potansiyel tehditleri daha iyi anlamak için yapay zeka araçları kullanmaktadır. Bu araçlar, uzmanlara, yüksek doğruluk oranlarıyla anomalilerin erken tespiti için matematiksel ve olasılıksal temelli modellerle, verileri otomatik bir şekilde tarama yeteneği sağlamaktadır (CISA, 2024b).

- **Gelişmiş Ağ Anomali Uyarısı (*Advanced Network Anomaly Alerting*)**

  Tehdit avcılığı ve Güvenlik Operasyon Merkezi (*Threat hunting and Security Operations Center*-SOC) analistlerine Ulusal Siber Güvenlik Koruma Sistemi (*National Cybersecurity Protection System*-NCPS) tarafından günde

terabaytlarca veri sağlanmaktadır. Yeterli eğitim verisi ve zaman verildiğinde birçok ağ saldırısı makine öğrenmesi tarafından olasılıksal olarak belirlenebilmektedir. Alınan uyarıların daha da iyileştirilmesi ve toplanan bilgilere dayanan ve konu uzmanlığıyla desteklenen ek otomatik uyarılar üretmek için otomatik araçlar kullanılır. Bu araçlar, CISA analistlerine, anomalilerin zamanında ve doğru olarak tespit edilmesini sağlamak için matematiksel ve olasılıksal temelli modellerle ağ tarama yeteneği sağlamaktadır(CISA, 2024b).

- **Yapay Zeka Güvenliği ve Sağlamlığı (*AI Security and Robustness*)**
  Yapay Zeka teknolojilerinin edinimi, geliştirilmesi, dağıtımı ve bakımını yönetmek için geliştirilen çerçeveler, süreçler ve test araçları da CISA tarafından geliştirilmektedir. CISA bünyesindeki teknoloji geliştiricileri, yapay zeka sistemlerinin güvenilir, sağlam ve güvenli çalışmasını sağlamak için yapay zeka ile geliştirilmiş araçlar kullanmaktadır. Bu araçlardan, veri işlemeyi hızlandırarak kurum içinde yapay zeka teknolojisinin kullanımını geliştirmek için makine öğrenimi ve doğal dil işleme gibi yöntemlerde faydalanılmaktadır(CISA, 2024b).
- **Kritik Altyapı Anomali Uyarısı (*Critical Infrastructure Anomaly Alerting*)**
  CISA tarafından geliştirilen *Cyber Sentry* programı ile kritik altyapı ağları izlenmektedir. Program kapsamında, tehdit analistleri, endüstriyel kontrol sistemleri ile gözetleyici kontrol ve veri toplama sistemleri (ICS/SCADA) dahil olmak üzere bilgi teknolojileri ve operasyonel teknoloji ağlarındaki siber ve fiziksel verileri incelemek için geliştirilecek anomali algılama ve makine öğrenimi yeteneklerine ihtiyaç duyulmuştur. Kritik Altyapı Anomali Uyarı modeli, bu bilgilerin işlenmesinde ihtiyaç duyulan yapay zeka yardımını sağlamaktadır(CISA, 2024b).
- **Siber Tehdit İstihbarat Akışı Korelasyonu (*Cyber Threat Intelligence Feed Correlation)***
  Siber Tehdit İstihbaratı Akış Korelasyonu, birden fazla gelen bilgi akışı arasında korelasyon sağlamak için yapay zeka yeteneklerini kullanmaktadır. Oluşturulan yapay zeka, algoritmanın, görevi yerine getirmenin en verimli

yollarını öğrenmek için verileri ve sonuçları kullanmasına olanak tanımaktadır. Sistem ayrıca, tehdit aktörlerinin taktik, teknik ve prosedürlerinin sürekli gözetimini sağlamak için özelleştirilebilmektedir(CISA, 2024b).

- **Kötü Amaçlı Yazılımların Tersine Mühendisliği (*Malware Reverse Engineering*)**

  Kötü amaçlı yazılımların tersine mühendisliği ve genel olarak yazılım analizi, siber güvenlik alanında kritik önem taşımaktadır. CISA tarafından uygulanan Tehdit Odaklı Tersine Mühendislik (*Threat Focused Reverse Engineering*-TFRE), ileri mühendislik ve derin öğrenme tekniklerinden yararlanarak daha verimli siber tehdit istihbaratı üretilmesini sağlamaktadır. Bu araçlar ile taktikleri, teknikleri ve prosedürleri açığa çıkararak kötü niyetli aktörlerin geliştirme yaşam döngüsünün[1] anlaşılması sağlanmaktadır(CISA, 2024b).

- **Operasyonel Faaliyetler Kaşifi (*Operational Activities Explorer*)**

  CISA'nın Operasyon Merkezindeki görevliler ve analistler, devam eden operasyonel faaliyetleri izlemek için yapay zeka destekli bir gösterge paneli kullanmaktadır. Bu yapay zeka, ulusal kritik sektörler üzerindeki potansiyel etkileri değerlendirerek diğer devlet kurumları ve kritik altyapı operatörleri ile iş birliği yapılabilecek eylem planları ve stratejiler önermektedir. Bu öneriler, geçmiş siber güvenlik ve altyapı güvenliği bilgileri ile önceki operasyonel müdahale faaliyetlerinin yanı sıra gerçek zamanlı yeni olay verileri (açık kaynak raporları, ortak raporlar ve siber güvenlik sensörlerinden alınan veriler) kullanılarak yapılmaktadır.

ABD'de ayrıca Ulusal Güvenlik Ajansı (*National Security Agency*-NSA) tarafından makine öğrenmesinin siber alanda kullanıldığı bilinmektedir. Bununla birlikte ABD özel sektörü yapay zeka siber güvenlik çözümlerinde önde gelen ülkelerden biri olarak karşımıza çıkmaktadır. Bu siber güvenlik çözümlerinde bazıları;

[1]Yazılım projelerinde izlenen süreci tanımlamak için kullanılmaktadır. Yazılımın planlanmasından bakımına kadar olan tüm adımları kapsamaktadır.

- **AI2**

  Massachusetts Teknoloji Enstitüsü (*Massachusetts Institute of Technology*-MIT) Bilgisayar Bilimi ve Yapay Zeka Laboratuvarı (*Computer Science and Artificial Intelligence Laboratory*-CSAIL) ve PatternEx[2], uzmanlardan gelen girdileri birleştirerek siber saldırıları tahmin etmek için AI2 isimli bir yapay zeka modeli geliştirmişlerdir(Veeramachaneni vd., 2016).

  Güvenlik analistlerinin deneyimlerini son teknoloji makine öğrenmesi ile birleştirerek analist-döngü-içi bir güvenlik sistemi sunulmaktadır. Bu sistem, dört ana bileşenden oluşmaktadır: Büyük Veri Davranış Analitiği Platformu, çeşitli aykırı değer tespit yöntemlerinden oluşan bir takım, güvenlik analistlerinden geri bildirim toplama mekanizması ve denetimli öğrenme modülü. Bu bileşenler bir arada çalıştığında, tespit oranı ortalama 3,41 kat artmakta ve yanlış pozitifler beş kat azalmaktadır(Veeramachaneni vd., 2016). AI2 sistemi, ham veriden farklı varlıkların davranışlarını hesaplayan bir büyük veri işleme sistemi olarak yoğunluk, matris ayrıştırma ve en önemlisi çoğaltıcı yapay sinir ağları (AE) gibi yöntemlerle sistemlerdeki anomalileri tespit etmektedir. Sistem, analistlerin belirlediği anomali değerlerini toplayarak bu veriler ile denetimli öğrenme modülünü besleyen bir geri bildirim mekanizmasını ve analist geri bildirimlerine dayanarak yeni olayların normal mi yoksa kötü niyetli mi olduğunu tahmin eden modeller öğrenen bir denetimli öğrenme modülünü içermektedir(Veeramachaneni vd., 2016).

  AI2 sistemi, güvenlik analistlerinin tecrübesini makine öğrenimi teknikleriyle birleştirerek yeni saldırıları tespit etmekte ve saldırıya karşı gerekli önlemlerin alınması için zaman kazandırmaktadır. Sistem, makine öğrenmesi tabanlı olduğu için zamanla analist geri bildirimlerini toplayıp tespit oranını artırarak yanlış pozitif oranını azaltacak, bu da verimliliği artırıp gereksiz alarm sayısını azaltacaktır(Veeramachaneni vd., 2016).

[2] Makine öğrenmesi ve yapay zeka teknolojileri kullanarak siber güvenlik tehditlerini tespit eden ve önleyen uygulamalar geliştirilmesi için kurulmuş bir MIT iştiraki.

- **Amazon Macie**

  Amazon Macie, 2017 makine öğrenmesinin örüntü eşleştirme yeteneklerini kullanan bir bilgi güvenliği hizmetidir. Yapay zeka, Macie'nin *Amazon Web Services*[3] (AWS) üzerindeki hassas verileri bulması, sınıflandırması ve koruması için araçlar sağlamaktadır(Amazon, 2024a).

  Amazon Macie, çeşitli veri türlerini ve içeriklerini yorumlamak ve sınıflandırmak için doğal dil işleme[4] (*Natural Language Processing*-NLP) yöntemlerini kullanmaktadır. NLP, bilgisayar bilimi ve hesaplamalı dil bilimi ilkelerini kullanarak ve bilgisayarlar ile insan dili arasındaki etkileşimleri inceleyerek bilgisayarların dil verilerini anlaması ve çözmesi üzerine çalışmalar yürütülen alandır. Amazon Macie, inceleme sonucunda bulduğu bulguları önceliklendirmekte ve değerlendirilen verilere otomatik olarak risk puanı atamaktadır. Bu süreç, son kullanıcıların en kritik uyarılara öncelik vermesine yardımcı olur(Sipola vd., 2023).

- **Deep Instinct**

  Deep Instinct, uçtan uca derin öğrenmeyi siber güvenlik alanında uygulayan ilk şirketlerden biridir(Deep Instinct, 2024a). Şirket, geliştirdiği çözümler ile derin öğrenme algoritmalarının kullanılarak kötü amaçlı yazılımlarda bulunan yapıların tanımlanmasını mümkün kılmıştır. Deep Instinct tarafından geliştirilen uygulamalar, bir organizasyonun tüm seviyelerindeki kötü amaçlı yazılımlarının tespit edilip çalıştırılmalarının önlenmesini amaçlamaktadır (Sipola vd., 2023). Deep Instinct tarafından geliştirilen yapay zeka ve sinir ağı

[3]Amazon'un bulut bilişim hizmetleri sunan platformudur.

[4]Bilgisayarların insan dilini anlama, yorumlama ve üretme yeteneklerini geliştirmeye odaklanan bir yapay zeka alanıdır. Dil bilimi, bilgisayar bilimi ve yapay zeka tekniklerini kullanarak metin ve konuşma verilerinin işlenmesi çalışılmaktadır.

teknolojileri, Spora[5], Wannacry[6], NotPetya[7] ve BadRabbit[8] gibi siber saldırılara karşı, belirli bir siber güvenlik merkezinden yardım almadan ve hedef sistemlere kötü amaçlı yazılımlar sızmadan tespit etmek için kullanılmaktadır (Deep Instinct, 2024b).

Deep Instinct tarafından yapay sinir ağları kullanılarak geliştirilmiş DIANNA isimli bir siber güvenlik asistanı da bulunmaktadır. DIANNA, zararlı dosyalar, betikler[9] ve belgeler gibi tehdit dosyalarını analiz etmekte ve bu dosyaların içeriğini doğal dile çevirerek kötü amaçlı yönlerini açıklamaktadır. Bilinmeyen tehditler hakkında bilgi sağlanması ve sıfırıncı gün[10] saldırılarının anlaşılmasının kolaylaştırılmasına yardımcı olmaktadır. Ayrıca iş akışlarını optimize ederek, rutin görevleri otomatikleştirerek ve hızlı olay

---

[5]Karmaşık bir fidye yazılımıdır. Diğer fidye yazılımlarından farklı olarak, dosyaların şifrelenmesi sırasında güçlü bir şifreleme yöntemi kullanır ve kullanıcıları fidye ödemeye ikna etmek için sosyal mühendislik taktikleri uygulanır. Fidye ödemesi yapıldıktan sonra dosyaların kurtarılmasını sağlayan bir sisteme sahiptir. Bu sistemin çalışması mağdur olanların fidye ödeme ihtimalini arttırmaktadır(Meskauskas, 2024).

[6]Mayıs 2017'de büyük bir küresel siber saldırıya neden olan bir fidye yazılımıdır. Microsoft Windows işletim sistemlerindeki bir güvenlik açığını hedefleyen bu yazılım, kullanıcıların dosyalarını şifreleyerek erişilemez hale getirir ve şifreyi çözmek için fidye ödenmesini talep eder. Saldırı, dünya genelinde birçok kurum ve kuruluşu etkileyerek büyük zararlar vermiştir(Cloudflare, 2024a).

[7]Haziran 2017'de yayılan ve büyük etkileri olan bir zararlı yazılımdır. Başlangıçta Petya olarak bilinen bir fidye yazılımının varyantı olduğu düşünülmüştür. Ancak, NotPetya'nın asıl amacı dosyaları geri döndürülemez şekilde silmek olduğundan, klasik bir fidye yazılımı olmaktan ziyade bir wiper (silici) yazılımı olarak kabul edilmektedir. Ukrayna'da başlayıp hızlıca yayılan bu yazılım, birçok global şirketin operasyonlarını ciddi şekilde aksatmıştır(Cloudflare, 2024b).

[8]Ekim 2017'de ortaya çıkan bir fidye yazılımıdır. Öncelikle Rusya ve Doğu Avrupa'yı hedef almıştır. BadRabbit, sahte bir Adobe Flash güncellemesi gibi görünen bir sosyal mühendislik saldırısı yoluyla yayılır. Kullanıcılar bu sahte güncellemeyi indirip kurduklarında, yazılım dosyalarını şifreleyerek fidye talep etmektedir(Blackberry, 2024).

[9]Belirli bir görevi veya işlemi otomatikleştirmek için yazılmış bir dizi talimattır. Genellikle bir programlama veya komut dosyası dilinde yazılır ve bilgisayar tarafından yürütülür.

[10]Siber güvenlikte, bir yazılımda keşfedilen ve geliştiricinin haberi olmadan saldırganlar tarafından aktif olarak istismar edilen güvenlik açığını ifade eder. Bu tür açıklar, geliştirici tarafından keşfedilip bir yama veya düzeltme yayımlanana kadar korunmasız kalır.

önceliklendirme sağlayarak siber güvenlik ekiplerine destek olmaktadır. Bu sanal yapay zeka yardımcısı, Deep Instinct'in diğer derin öğrenme ürünleri ve yetenekleriyle entegre olarak, güvenlik geliştirmeyi ve yanlış pozitifleri azaltmayı amaçlamaktadır(Deep Instict, 2024).

- **SparkCognition DeepArmor**

  SparkCognition, 2013 yılında kurulan ve merkezi Austin, Teksas'ta bulunan bir yapay zeka teknoloji şirketidir. Şirketin DeepArmor isminde bulut destekli bir yapay zeka ürünü bulunmaktadır. DeepArmor, kapsamlı siber güvenlik özellikleri sunan, istemci tarafı öğeleri, bulut barındırma yönetim konsolu ve küresel bulut hizmetlerini birleştiren SaaS[11] tabanlı bir koruma platformudur. DeepArmor'un temel unsurları, Uç Nokta Koruma Ajanı ve Uç Nokta Koruma Bulutu'dur. Ajan, sistem izleme, yapay zeka tabanlı tehdit algılama ve uç nokta koruma sağlamakta, gerçek zamanlı olarak süreçleri, dosyaları ve komut dosyalarını izlemekte ve kötü niyetli faaliyetleri tespit etmek ve tahmin etmek için makine öğrenmesi modellerini kullanmaktadır. *Google Cloud Platform*[12] (GCP) üzerinde çalışan bulut barındırma hizmetleri için, ölçeklenebilirlik, erişilebilirlik ve güvenlik sunmaktadır. Şirkete göre bu yapı, operasyonel maliyetleri azaltmakta, yüksek erişilebilirlik sağlamakta ve şifreleme, kimlik doğrulama ve denetim kaydı gibi özellikler sayesinde bakım süreçlerini kolaylaştırmaktadır(SparkCognition, 2018).

- **Vectra AI**

  Vectra AI, 2012 yılında kurulan ve merkezi San Jose, Kaliforniya'da bulunan bir siber güvenlik şirketidir. Şirket, hibrit saldırıların tespiti, soruşturulması ve müdahale edilmesi için yapay zeka ürünleri geliştirmektedir. Ana ürünleri Vectra AI platformudur(VectraAI, 2024).

  Vectra AI platformu, siber güvenlik alanında tehdit tespiti ve yanıt verme yetenekleri sunan bir çözümdür. Makine öğrenmesi ve yapay zeka teknolojilerini kullanarak ağ trafiğini sürekli olarak analiz etmekte ve

---

[11]Hizmet olarak yazılım (*Software as a Service*, SaaS), kullanıcıların internet üzerinden yazılım uygulamalarına erişim sağlamasına olanak tanıyan bir bulut bilişim modelidir.

[12]Google'ın bulut bilişim hizmeti sunan platformudur.

anomallileri ya da aykırılıkları tespit etmektedir. Bu sayede, ağda gerçekleşen saldırıları, veri ihlallerini ve diğer tehditler hızla belirlenip müdahale edebilir. Vectra AI, kullanıcı davranışlarını izleyerek, ağ içi hareketleri analiz ederek ve saldırganların bıraktıkları izleri takip ederek, siber güvenlik ekiplerinin tehditlere karşı önleyici bir savunma oluşturmasına yardımcı olmaktadır(VectraAI, 2024). Platform, ayrıca, bulut tabanlı hizmetlerle de entegre çalışabilmektedir. Vectra AI platformu, tehdit tespiti ve yanıt verme yeteneklerini geliştirmek makine öğrenmesi yöntemlerinden faydalanmaktadır. Hem denetimli hem de denetimsiz öğrenme tekniklerini kullanarak ağ trafiğini analiz etmekte ve potansiyel tehditleri belirlemektedir(VectraAI, 2024).

Denetimli öğrenme, insan analistlerin belirli tehditleri ve normal davranışları tanımladığı etiketli verilerle algoritmaların eğitilmesidir. Bu, sistemin kimlik bilgisi kötüye kullanımı, iç keşif ve veri sızdırma gibi kötü niyetli faaliyetlerle ilişkilendirilen kalıpların algoritmaya tanıtılmasına yardımcı olmaktadır. Denetimsiz öğrenme ise, önceden tanımlanmış etiketler olmadan ağ trafiğini analiz ederek anormallikleri tespit etmeyi amaçlamaktadır. Bu yöntem, daha önce bilinmeyen veya yeni ortaya çıkan tehditleri belirlemede etkilidir(VectraAI, 2024).

Bu yöntemler sayesinde, Vectra AI platformundaki makine öğrenmesi modelleri, konuşlandırıldıkları ortamda sürekli olarak öğrenip uyum sağlamaktadır. Büyük miktarda ağ meta verisi[13], sistem günlükleri ve diğer bağlamsal verileri işleyerek doğru tehdit tespiti sağlamayı amaçlamaktadır(VectraAI, 2024).

- **Palo Alto Networks Cortex XDR**

  Palo Alto Networks, siber güvenlik çözümleri sağlayan bir şirkettir. Merkezi Kaliforniya'da bulunan şirket 2005 yılında kurulmuştur. Cortex XDR, Palo

[13]Meta veri (*metadata*), bir veri kümesi hakkında bilgi sağlayan verilerdir. Meta veri, genellikle ana verinin özelliklerini, yapısını ve bağlamını açıklamak için kullanılır. Örneğin, bir dosyanın meta verileri, dosya adı, boyutu, oluşturulma tarihi, son değiştirilme tarihi ve dosya türü gibi bilgileri içerebilir.

Alto Networks tarafından geliştirilen, yapay zeka ve makine öğrenmesi teknolojileriyle güçlendirilmiş bir kapsamlı algılama ve yanıt verme (*Extended Detection and Response*-XDR) platformudur. Bu platform, kurumsal güvenlik ekiplerine, gelişmiş tehditleri tespit etmek ve önlemek için gerekli olan geniş kapsamlı veri analizi ve işleyişe dair bilgiler sağlamaktadır. Cortex XDR, uç nokta, ağ, bulut ve kimlik verilerini entegre ederek, makine öğrenmesi algoritmaları ile sürekli olarak kullanıcı ve ağ davranışlarının profilini çıkartmaktadır. Bu sayede, anormal aktiviteleri ve saldırı belirtilerini tespit ederek, saldırıların gerçekleşmeden önce önlenmesini mümkün kılmayı amaçlamaktadır. Cortex XDR AI destekli algoritmalar ile güvenlik ekiplerinin dikkatini çekmesi gereken önemli olayları önceliklendirerek, buna bağlı olarak sistem olay incelemelerinin hızlandırılmasını ve tehditlere karşı hızlı yanıt verilmesini sağlamaktadır(PaloAltoNetworks, 2023).

- **Zscaler Internet Access**

  Zscaler, bulut tabanlı bilgi güvenliği çözümleri sağlayan bir şirkettir. 2008 yılında Kaliforniya'da kurulmuştur. *Zscaler Internet Access* (ZIA), tüm kullanıcılar, uygulamalar ve cihazlar için yapay zeka destekli koruma sağlamak amacıyla geliştirilmiş bir üründür. ZIA, siber güvenliği buluta taşıyarak ve sıfır güven[14] (*zero trust*) mimarisi ile eski güvenlik altyapılarının yerini almayı hedeflemektedir. Yapay zeka destekli tehdit analizi ve kapsamlı veri koruma özellikleri ile fidye yazılımlarına ve kötü amaçlı yazılımlara karşı koruma sağlamaktadır(Zscaler, 2024).

3.1.2 Birleşik Krallık

Birleşik Krallıkta kamu sektöründe siber güvenlik alanında gerçekleştirilen teknik çalışmalar Ulusal Siber Güvenlik Merkezi (*National Cyber Security Centre* - NCSC) tarafından gerçekleştirilmektedir. NCSC Birleşik Krallık'ın siber tehditler ve Bilgi Güvencesi konularında ulusal teknik otoritesidir(United Kingdom, 2024). Hükümet

[14]Geleneksel güvenlik yaklaşımlarının aksine, ağa dahil olan hiçbir kullanıcı veya cihazın güvenilir kabul edilmediği bir siber güvenlik çerçevesidir.

İletişim Karargahı[15] (Government Communications Headquarters - GCHQ) altında Ekim 2016'da kurulan NCSC'nin merkezi Londra'da bulunmaktadır. GCHQ, *CERT-UK* ve diğer kurumların siber güvenlik yapılanmalarını uzmanlıklarını bir araya getirmiştir(NCSC, 2024).

NCSC, KOBİ'ler, büyük kuruluşlar, devlet kurumları, vatandaşlar ve diğer kurumlar için koordinasyon sağlamaktadır. Ayrıca, kolluk kuvvetleri, savunma, Birleşik Krallık'ın istihbarat ve güvenlik ajansları ile uluslararası ortaklarla iş birliği içerisinde çalışmaktadır(NCSC, 2024).

NCSC tarafından hizmete sunulan projelerin ve servislerin oluşturduğu sistemin bütünü Aktif Siber Savunma (*Active Cyber Defence* - ACD) olarak adlandırılmaktadır. Bu sistemin sağladığı servisler içerisinde makine öğrenmesinden yararlananlar da bulunmaktadır. Ayrıca ACD kapsamında, yapay zeka alanında araştırma geliştirme faaliyetleri gerçekleştirilmektedir (Levy, 2019; NCSC, 2020, 2022, 2023).

Birleşik Krallık'ın veri bilimi ve yapay zeka ulusal enstitüsü olan Alan Turing Enstitüsü ile birlikte ACD'ye dahil olan *Mail Check*[16] servisinin verileri ile bir yapay zeka eğitilmesi ve sisteme entegre edilmesi konusunda çalışmalar yürütülmektedir(Levy, 2019). Buna ek olarak makine öğrenmesi yardımıyla algoritmik olarak üretilen alan adlarının (*Algorithmically Generated Domains*-AGD) tespitine yönelik çalışmalar yapılmıştır(NCSC, 2020). Bu alan adları, kötü niyetli aktörler tarafından C2[17] sunucuları ile buluşma noktası olarak kullanılmaktadır(NCSC, 2020). Bu proje ile kelime tabanlı AGD'ler ile üç farklı zararlı yazılım türünün

[15]Birleşik Krallık hükümeti ve silahlı kuvvetlerine sinyal istihbaratı ve bilgi güvencesi sağlamakla sorumlu bir istihbarat ve güvenlik kuruluşudur(GCHQ, 2024).

[16]Kuruluşların e-posta güvenliği uyumluluğunu değerlendirmeyi ve e-posta alan adlarının sahtecilik yoluyla kötüye kullanılmasını önlemeyi amaçlayan, güvenli e-posta standartları oluşturulmasına yardımcı bir servistir.

[17]Komuta ve Kontrol (*Command and Control*) sunucusu genellikle kötü amaçlı yazılımların veya botnetlerin, siber saldırganların kontrolü altında olan merkezi bir sunucuya bağlanarak komutlar alması ve veriler göndermesi için kullanılır.

tanımlanması için makine öğrenimi ve doğal dil işleme teknikleri kullanılmıştır (NCSC, 2020).

Bristol Üniversitesi ile veri analitiği ve makine öğrenmesini kullanılarak *BGP hijacking*[18] verileri üzerinde çalışmalar yürütülmektedir. Çalışmanın amacı yanlış pozitif verileri daha verimli tespit etmek ve kaçırılma olasılığı yüksek görünen rotaları öngörmeye çalışmaktır(NCSC, 2022). Ayrıca, kötü niyetli olduğu Doğrulanmış web sitelerinin veri seti kullanılarak eğitilen bir makine öğrenmesi algoritması halihazırda veri setinin içinde bulunmayan zararlı siteleri tespit etmek için eğitilmektedir(NCSC, 2023).

Birleşik Krallık'ta da güçlü bir siber güvenlik sektörü bulunmakla birlikte yapay zeka ve makine öğrenmesi konusunda Darktrace öne çıkmaktadır. Darktrace Birleşik Krallık Cambridge'de 2013 yılında kurulmuştur. Ana ürünü *Darktrace ActiveAI Security Platform* olan Şirket siber güvenlik sektöründe faaliyet göstermektedir.

Darktrace ActiveAI Güvenlik Platformu, günümüz siber güvenlik dünyasında karşılaşılan zorlukları aşmak için tasarlanmış kapsamlı bir çözümdür. Platform, yeni tehditlerle başa çıkmak ve kurumlara bütünsel bir siber güvenlik çözümü sunmak amacıyla geliştirilmiştir.

Platform, kendini sürekli olarak işletmenin değişen verilerinden öğrenen ve dış tehdit istihbaratıyla zenginleştirilen bir yapay zekâ modeline dayanmaktadır(Darktrace, 2024). Bu model, yalnızca bilinen tehdit verilerine dayalı algılama yapmanın yanı sıra, işletmeye özgü anormallikleri belirleyerek daha önce görülmemiş tehditleri de tespit

[18]BGP (*Border Gateway Protocol*) hijacking, internetin yönlendirme sisteminde ciddi bir güvenlik sorunudur. BGP, internet üzerindeki farklı otonom sistemler arasında yönlendirme bilgilerini değiş tokuş etmek için kullanılan bir protokoldür. *BGP hijacking*, kötü niyetli bir aktörün yanlış yönlendirme bilgileri yayarak, trafik akışını kontrol etmesi veya başka bir yöne yönlendirmesi durumudur. Bu, internet trafiğinin yanlış bir yöne gitmesine, kesintilere, veri kaybına veya gizlilik ihlallerine yol açabilir.

etmektedir. Bununla birlikte platform, güvenlik operasyonlarını dönüştürerek, elle yapılan önceliklendirme sürecini ortadan kaldırmayı hedefleyen ve ilgili tüm olayları otomatik olarak inceleyen bir AI analisti sunmaktadır.

Platform, ağ, operasyonel teknolojiler, bulut ve e-posta gibi çeşitli alanlarda veri toplayarak ve işletmenin benzersiz verilerini analiz ederek gerçek zamanlı öğrenme gerçekleştirmektedir(Darktrace, 2024). Ayrıca sunduğu tehdit yönetimi ile zayıf noktaları ve saldırı yollarını önceden belirlemeyi amaçlamakta ve gerektiğinde eğitim simülasyonları ve phishing testleri ile insan faktörünü güçlendirmeye yönelik çalışmalar yapılabilmesini de sağlamaktadır(Darktrace, 2024). Bu özellikleri ile platform, sürekli öğrenen yapay zekâ modeli ve otonom yanıt yeteneklerini birleştirerek işletmelerin güvenlik süreçlerini dönüştürmeyi ve işletmelerin dijital varlıklarını koruma altına almayı hedeflemektedir (Darktrace, 2024).

Birleşik Krallıkta NCSC tarafından yürütülen bilinçlendirme faaliyetlerine rağmen KOBİ'ler arasında makine öğrenmesi ve siber güvenlik alanındaki farkındalık hala istenilen noktada bulunmamaktadır(Rawindran vd., 2021). KOBİ'lerin temel siber güvenlik önlemlerini uygulamış olmalarına rağmen, makine öğrenmesi teknolojilerini anlama ve kullanma düzeyleri hala sınırlıdır(Rawindran vd., 2021). KOBİ'lerdeki yöneticiler, makine öğrenmesi ve yapay zekayı güvenlik çerçevelerine entegre etme konusunda daha fazla bilgi edinme ve siber güvenlik yeteneklerini geliştirme isteği göstermektedir(Rawindran vd., 2021).

3.1.3. Fransa

Fransa'da da makine öğrenmesinin siber güvenlik alanında kullanımı gelişmekte olan bir sektör olarak karşımıza çıkmaktadır. Kamu ve akademide bu konuda çalışmalar yapılması teşvik edilmektedir. 2009 yılında Fransa Başbakanı'na bağlı olarak kurulan Ulusal Bilgi Sistemleri Güvenliği Ajansı (*Agence nationale de la sécurité des systèmes d'information* - ANSSI), Fransa'nın Ulusal Siber Güvenlik Ajansı'dır ve Fransa'nın kritik bilgi sistemlerini korumak ve siber güvenliği teşvik etmekle görevlidir(ANSSI, 2024a). ANSSI'nin sorumlulukları arasında kritik altyapının korunması, CERT-FR

ekibi aracılığıyla olaylara müdahale koordinasyonu ve siber güvenlik standartlarının ve düzenlemelerinin geliştirilmesi yer almaktadır(ANSSI, 2024a). Ajans ayrıca siber güvenlik kültürünü teşvik etmek için farkındalık ve eğitim çalışmaları yürütmektedir. Siber güvenlik teknolojilerinde araştırma geliştirme faaliyetlerini desteklemekte ve küresel siber tehditlerle başa çıkmak için uluslararası iş birliğini sağlamaktadır(ANSSI, 2024a).

ANSSI yaptığı çalışmalara paralel olarak açık kaynak olarak geliştirdiği bir makine öğrenmesi projesini kullanıcılarla paylaşmaktadır. SecuML, Bilgisayar Güvenliğinde Makine Öğrenmesinin kullanımını teşvik etmeyi amaçlayan ve Python[19] programlama dili ile geliştirilmiş bir araçtır(ANSSI, 2024b). GPL2+[20] lisansı altında dağıtılmaktadır. Siber güvenlik uzmanlarının tespit ve algılama modellerini kolayca eğitmesine olanak tanır ve sonuçları görselleştirmek ve modellerle etkileşim kurmak için bir web kullanıcı arayüzü sağlamaktadır(ANSSI, 2024b).

Fransa'da yapay zeka ve makine öğrenmesinin siber güvenlik alanına uygulanması üzerine çalışan bir başka kurum ise Ulusal Dijital Bilimler ve Teknolojiler Araştırma Enstitüsü'dür (*Institut national de recherche en sciences et technologies du numérique - Inria*). Enstitü siber güvenlik alanında pek çok çalışma yürütmekte ve makaleler yayınlamaktadır(Inria, 2024b). Bu projeler arasında makine öğrenmesi ve siber güvenliği birleştiren, *Machine Learning Network System Security* (MLNS2) projesi karşımıza çıkmaktadır. MLNS2 projesi, simbox[21] sahtekarlıkları ve kötü amaçlı yazılım yayılımı ile mücadele etmek için Makine Öğrenmesi, Ağ, Sistem ve Güvenlik gibi birçok disiplinden yararlanarak bir sistem tasarlanmasını ve konu üzerine araştırma yapılmasını amaçlamaktadır(Inria, 2024a).

---

[19]Python, genel amaçlı programlama dilleri arasında popüler olan, basit sözdizimi ve geniş kütüphane desteği ile kullanıcı dostu ve güçlü bir dildir.

[20]Yazılımın GNU Genel Kamu Lisansı sürüm 2 veya daha sonraki bir sürümü altında dağıtılabileceğini belirten bir lisanslama terimidir.

[21]Simbox sahtekarlığı, uluslararası çağrıların yerel çağrılar gibi görünmesini sağlayarak telekom operatörlerinin gelir kaybına yol açan bir dolandırıcılık türüdür.

Fransa, siber güvenlik konusunda güçlü bir özel sektöre ve makine öğrenmesinde faydalanılan ürünlere de sahiptir;

- **Thales Guavus-IQ**

  Thales Group, merkezi Fransa'nın Paris kentinde bulunan, savunma, havacılık, uzay, ulaşım, güvenlik ve dijital kimlik sektörlerinde faaliyet gösteren çok uluslu bir şirkettir. Şirketin veri güvenliği ve siber güvenlik alanında da çözümleri bulunmaktadır(Thales, 2024).

  Thales şirketi olan Guavus, endüstriyel IoT, havacılık, ulaşım ve dijital güvenlik uygulamaları için çözümler sunmaktadır. Guavus-IQ, bulut üzerinde analitik, yapay zeka ve makine öğrenimi tekniklerini kullanarak 5G Mobil Ağ Operatörlerinin güvenlik seviyesini, iş sürekliliğini, işletim verimliliğini artırmayı, 5G güvenlik gereksinimlerini karşılamayı ve kullanıcılar, makineler ve IoT cihazları için daha kaliteli hizmet sağlanmasını hedeflemektedir. Guavus tarafından, telekom alanı deneyimi, veri bilimi ve yazılım mühendisliği birleştirilerek dünyanın önde gelen ağ operatörlerinin her gün petabaytlarca veri toplamak ve analiz etmek için kullandığı bir sistem oluşturulmuştur(Amazon, 2024b).

- **Airbus Defence & Space**

  Airbus, dünya genelinde ticari ve askeri uçak, savunma ve uzay endüstrilerinde faaliyet gösteren çok uluslu bir havacılık şirketidir. 1970 yılında kurulmuş olup merkezi Fransa'nın Toulouse kentinde bulunmaktadır(Airbus, 2024). Airbus, siber güvenlik çözümleri sunmak üzere Airbus CyberSecurity adında bir yan kuruluş kurmuştur. Bu kuruluş, kritik altyapıları, askeri sistemleri, kamu kurumlarını ve özel sektör kuruluşlarını siber tehditlere karşı korumak amacıyla çözümler geliştirmektedir(Airbus CyberSecurity, 2024a). Şirket tarafından *Cognification*[22] olarak sınıflandırılan yapay zeka ve makine öğrenmesi destekli çözümler bulunmaktadır. Şirket, yapay zekayı tehditleri

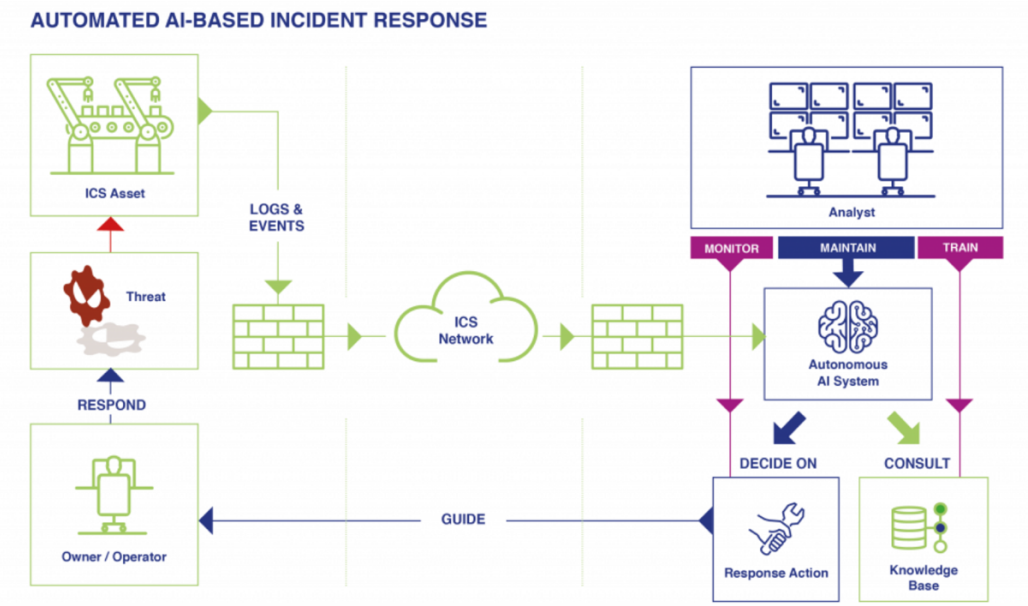
[22]Bilişselleştirme - şirket tarafından yapılan tanıma göre nesneleri ve sistemleri yapay zeka ve makine öğrenmesi kullanarak daha akıllı hale getirme süreci(Airbus CyberSecurity, 2024b).

tanımlamak, saldırıları tespit etmek ve yanıtları otomatikleştirmek için kullanmaktadır. Airbus, yapay zeka destekli yeni nesil siber savunma araçları üretmekte ve siber analistlerini yapay zeka ile birlikte verimli çalışacak şekilde eğitmektedir (Airbus CyberSecurity, 2024b).

Şirketin üzerinde çalıştığı yapay zeka destekli projelerden biri Güvenlik Operasyon Merkezlerinin (*Security Operation Centre* - SOC) otomatik hale getirilmesi ile ilgilidir. SOC kurumlar içinde kurulmuş ağı siber olaylara karşı takip eden ve gerektiğinde müdahalede bulunan merkezlerdir. Otomasyon sistemlerinin, SOC'ta görev alan analistlerin karar verme sürecini desteklemesi ve onların yerini alarak maliyetleri azaltması hedeflenmiştir (Airbus CyberSecurity, 2024c). Ancak, bu durum analistlerin rolünü tamamen gereksiz kılmayacaktır. Bunun yerine, analistlerin çalışmaları ve odakları, tüm kararları kendileri vermekten, kararların büyük kısmını ele alan yapay zekayı izlemeye kayacaktır. AI sisteminin eğitim aşamasında, analistler tüm kararlarını gözden geçirerek ve doğrulukları ile başarısızlık durumunda sorumlu parametreler hakkında geri bildirim sağlayarak makine öğrenmesini destekleyeceklerdir.

Zaman ve veri ile, yapay zekanın bilgi tabanı büyüyecek ve sonuç olarak yanlış kararların sayısı azalacaktır. Bu operasyonel aşamada, analistler AI kararlarının örnek setlerini sürekli olarak gözden geçirebilir ve gerektiğinde olay müdahalesini devralabilir ve işleri her zamanki gibi yürütebilir. "Bu bağlamda, geleceğin analist rolü farklı olmayacak sadece makine öğrenimi bilgisi gibi ek beceri ve sorumluluklar gerektirecektir." (Airbus CyberSecurity, 2024c).

Şekil: 2.1. Otomatik Yapay Zeka Temelli Olay Müdahale sistemi



Kaynak: (Airbus CyberSecurity, 2024c)

- **Orange Cyberdefence**

  Orange, 1988 yılında kurulmuş Fransa merkezli bir GSM operatörüdür. Orange Cyberdefence ise şirketin siber güvenlik çözümleri sunan şirketidir. Orange Cyberdefence'in Security Event Intelligence çözümü, gelişmiş tehdit algılama, gerçek zamanlı analiz ve otomatik yanıt yetenekleri aracılığıyla siber güvenliği desteklemek için yapay zeka ve makine öğrenimini kullanmaktadır. Güvenlik bilgi ve olay yönetimi sistemleriyle entegre olarak kapsamlı güvenlik izleme, olay tespiti ve tehdit istihbaratı sağlamak için geliştirilmiştir. Bu çözüm, kuruluşların risklerini, olaylara müdahale sürelerini azaltmasına ve standartlara uyum sağlamasına yardımcı olmaktadır. Security Event Intelligence, potansiyel tehditleri zarar vermeden önce tespit etmek ve etkisiz hale getirmek için tahmine dayalı analitik ve davranışsal analiz kaynaklı makine öğrenmesi algoritmalarından yararlanmaktadır (Orange, 2024).

3.1.4. Türkiye

Ordinaryüs Profesör Doktor Cahit Arf Erzurum Atatürk Üniversitesinde 1959 yılında yaptığı "Makineler Düşünebilir mi ve Nasıl Düşünebilir?" başlıklı sunum ile Türkiye'de yapay zekaya ilişkin çalışmaların öncüsü olmuştur(Arf, 1959). Günümüze gelindiğinde ise ülkemizin yapay zeka teknolojilerine verdiği önemin bir göstergesi olan ve 20/08/2021 tarihli ve 31574 sayılı Resmî Gazetede yayınlanan 2021/18 sayılı Cumhurbaşkanlığı Genelgesi ile Ulusal Yapay Zekâ Stratejisi 2021-2025 yürürlüğe girmiştir. Bu sayede ülkemiz yapay zeka alanında strateji geliştiren sınırlı ülkeler arasında yerini almıştır. Stratejide ülkemizin refah seviyesini artmasına katkı sağlayacak ve yapay zeka teknolojisinin ülkemizdeki gelişimine destek olacak hedefler yer almaktadır. Bu hedefler altı stratejik amaç altında toplanmıştır(DDO, 2021).

1. Yapay Zeka Uzmanlarını Yetiştirmek ve Alanda İstihdamı Artırmak
2. Araştırma, Girişimcilik ve Yenilikçiliği Desteklemek
3. Kaliteli Veriye ve Teknik Altyapıya Erişim İmkânlarını Genişletmek
4. Sosyoekonomik Uyumu Hızlandıracak Düzenlemeleri Yapmak
5. Uluslararası İş Birliklerini Güçlendirmek
6. Yapısal ve İşgücü Dönüşümünü Hızlandırmak

Bu amaçlar doğrultusunda hem kamu hem özel sektörde yapay zekaya dair çalışmalar hızlanmıştır. Son olarak 28 Mayıs 2024 tarihli Milli Güvenlik Kurulu toplantısında yapay zeka hususu gündeme gelmiştir. Toplantı bildirisinin 7. maddesi:

> "*Yapay zeka alanında kaydedilen ilerlemenin, insanlık tarihinde yeni bir merhaleye geçilmesini mümkün kılabilecek büyük fırsatlar sunduğuna; bununla birlikte, bahse konu sahadaki potansiyelin birtakım sınamaları ve siber alanda oluşan yeni tehditleri de beraberinde getireceğine dikkat çekilmiştir. Türkiye'nin, yapay zeka çalışmalarının ilmi, askeri, iktisadi ve içtimai neticelerine hazırlıklı olmasının ve bu alanda ileri kabiliyetler geliştirmesinin önem ve önceliğine işaret edilmiştir*(MGK, 2024)."

olarak açıklanmıştır.

Açıklanan bildiriyle ülkemizin yapay zekanın siber alana olan etkisini de dikkate aldığı ve ilerleyen zamanlarda yapay zekanın Türkiye'nin her alanında öncelikli hale geleceği ve bu alana verilen önem anlaşılmaktadır.

Türkiye'deki siber güvenlik çalışmaları Bilgi Teknolojileri ve İletişim Kurumu (BTK) tarafından koordine edilmektedir. 5809 sayılı Elektronik Haberleşme Kanunu'nun 60'ıncı maddesinin 11,12 ve 13'üncü fıkraları çerçevesinde BTK'ya ulusal siber güvenliğin sağlanması hususunda çeşitli görevler verilmiştir. Siber güvenlik faaliyetleri BTK bünyesinde faaliyet göstermekte olan Ulusal Siber Olaylara Müdahale Merkezi (USOM) aracılığıyla yürütülmektedir. USOM 2013 yılı Mayıs ayında kurulmuş olup, 2016 yılından bu yana BTK bünyesinde yer almaktadır. USOM, ülke kapsamında yurtiçi ve yurtdışı kaynaklı siber tehditleri önlemek amacıyla alarm, uyarı ve duyurulara ilişkin faaliyetler yürütmekte, kritik durumlarda yerinde müdahale ekipleriyle olayın kontrolünün sağlanmasında önemli bir rol üstlenmektedir. Ayrıca 2013 yılından bu yana hazırlanan ve son olarak 2020-2023 yılları için yayınlanan Ulusal Siber Güvenlik Strateji ve Eylem Planı kapsamında çalışmalar yürütülmektedir.

Bu doğrultuda USOM, makine öğrenmesinden faydalanarak geliştirdiği AZAD isimli projeyi ülkemizin siber güvenliğinin sağlanmasında kullanmaktadır. AZAD bir *phishing*[23](oltalama) tespit sistemidir. Sistem önceden belirlenmiş ve oltalama olarak sınıflandırılmış alan adlarının listesini kullanarak öğrenen bir makine öğrenmesi modelinin sonuçları ve doğal dil işleme yöntemlerini birleştirerek yeni alan adlarının oltalama olup olmadığını tespit etmektedir.

AZAD RNN yönteminin bir alt modeli olan Uzun-Kısa Süreli Bellek (*Long Short-Term Memory* - LSTM) modeli ile eğitilmiştir. Geleneksel bir RNN, zaman içerisinde iletilen tek bir gizli katmana sahiptir, bu da ağın uzun vadeli örüntüleri öğrenmesini zorlaştırabilmektedir(Hochreiter & Schmidhuber, 1997). LSTM'ler, uzun süre bilgi tutabilen bir nöron geliştirilmesi ile bu sorunu çözmeyi hedeflemiştir(Hochreiter &

[23]Genellikle güvenilir hizmetler gibi gözükerek, kullanıcıların hassas bilgilerini (parola, kredi kartı numarası, telefon numarası vb.) ele geçirmeyi amaçlayan bir siber saldırı türüdür.

Schmidhuber, 1997). Bu nöron sayesinde LSTM yapay sinir ağları, verilerdeki uzun vadeli bağımlılıkları öğrenme yeteneğine sahiptir. Bu, dil çevirisi, konuşma tanıma ve seri tahminleme gibi görevler için LSTM'yi uygun hale getirmiştir. LSTM'ler ayrıca, görüntü ve video analizi için CNN'ler gibi diğer sinir ağı mimarileri ile kullanılabilir.

Tablo: 3.1. LSTM RNN Karşılaştırması

	LSTM	**RNN**
Bellek	Özel bir bellek nöronu vardır	Bellek nöronu yoktur
Eğitim yönü	Verileri hem ileri hem de geri yönde işlemek üzere eğitilebilir	Verileri yalnızca tek yönde işlemek üzere eğitilebilir
Eğitim maliyeti	Daha maliyetli	Daha az maliyetli
Uzun vadeli Örüntü öğrenme yeteneği	Yüksek	Düşük

Geçmiş verilerden faydalanılarak eğitilen AZAD test verilerinde yüzde doksa beş oranına yakın bir doğrulukta oltalama sitelerini tespit edebilmektedir. Siber güvenlik uzmanları tarafından tespit edilen hatalı sonuçlar tekrardan sistem tarafından eğitim verisi olarak kullanılmaktadır.

Türkiye'de özel sektörde de yapay zeka destekli siber güvenlik çalışmaları gerçekleştirilmektedir. Savunma Sanayi Başkanlığı (SSB) tarafından Savunma Sanayi Yapay Zeka Platformu (SSYZ) kurulmuştur. Platform, Savunma Sanayii Başkanlığı Stratejik Planı ve Ulusal Yapay Zekâ Stratejisi'ne uygun olarak yapay zekâ yeteneklerini savunma sanayiine sunmayı amaçlamaktadır. Bu platform, yüksek hacimli verisi olmayan, verilerini etiketleyemeyen, model eğitimi için zaman ve bütçe ayıramayan firmalar, kurumlar ve bireyler için uçtan uca bir yapay zekâ altyapısı sağlamayı amaçlamaktadır. Ayrıca, yüksek maliyetli donanım yatırımlarını ve mükerrer çalışmaları azaltarak, dışa bağımlılığı en aza indirmeyi

hedeflemektedir(SSYZ, 2024). Bu doğrultuda platform yarışmalar düzenlemekte, çalıştaylar organize etmekte ve belirli konularda projeler başlatarak sektörü davet etmektedir.

Siber güvenlik alanındaki yapay zeka çalışmaları yeni olmakla birlikte ülkemizde de çalışmalar sürmektedir;

- **DNSSense**

  Türkiye'de yapay zeka destekli bir siber güvenlik çözümü sunan bir şirkettir. Bu ürün, DDR 2.0, yapay zeka destekli bir DNS[24] Algılama ve Yanıt çözümüdür. Siber güvenliği otomasyon, gerçek zamanlı alan adı analizleri ve makine öğrenmesi kullanarak güçlendirmeyi hedeflemektedir. Yapay zeka ve Makine Öğrenmesi algoritmalarını kullanarak DNS tünelleme[25] gibi sofistike DNS tabanlı tehditlerin tespit edilmesini ve etkilerinin hafifletilmesini sağlamaktadır.

- **STM**

  Türkiye'nin önde gelen savunma şirketlerinden biri olan STM, 1991 yılında Savunma Sanayi İcra Komitesi (SSİK) kararıyla kurulmuştur. STM, siber güvenlik alanında yapay zeka kullanımı konusunda ARGE faaliyetlerine önem vermektedir. Bu kapsamda yapılan çalışmaların sonucu olarak siber açıklıklara karşı CyberAID isminde NVT üretebilen ve CVSS skor tahmini yapabilen bir yapay zeka çözümü geliştirilmiştir. Ayrıca CyberAID, eksik zafiyet bilgilerini tamamlamak ve benzer zafiyetleri ilişkilendirerek zafiyet tarama ve

---

[24] DNS (*Domain Name System*), internet üzerinde alan adlarını IP adreslerine dönüştüren bir sistemdir. Bu, kullanıcıların hatırlanması zor IP adresleri yerine daha kolay hatırlanabilen alan adları kullanarak web sitelerine erişmelerini sağlar. DNS, telefon rehberi gibi çalışır; bir alan adını girildiğinde, DNS sunucuları bu adı ilgili IP adresine çevirir ve böylece kullanıcıların doğru sunucuya yönlendirilmesini sağlanmaktadır.

[25] *DNS Tunneling*, verilerin DNS protokolü üzerinden gizlice iletilmesi için kullanılan bir tekniktir. DNS protokolünü kötüye kullanarak, veri transferi gerçekleştirir. Bu teknik, güvenlik duvarlarını ve diğer ağ güvenlik önlemlerini atlatmak için kullanılır ve genellikle siber saldırganlar tarafından veri sızdırma, komut kontrol iletişimi veya zararlı yazılımların çalıştırılması amacıyla kullanılmaktadır(PaloAltoNetworks, 2024).

değerlendirme çalışmalarına katkı sağlamak üzere geliştirilmiş modüller de içermektedir. Ayrıca, geliştirilen Pro-Post isimli bir başka modül aracılığıyla, yapay zeka kullanılarak zafiyetlerin potansiyel zararları ve saldırı sonrası alınabilecek önlemleri tahmin edilmektedir. Bunlarla birlikte Cyber Entity Recognition isimli bir modül ile veri kaynaklarından anahtar kelimeler seçilmesi ve sınıflandırılması sağlanmaktadır(STM, 2024).

3.2. Ülke İncelemelerine Göre Kullanım Alanları

3.2.1. Tehdit Tespiti İçin Makine Öğrenmesi

Güvenlik yaşam döngüsü üç aşamada incelenebilir; önleme, tespit ve tepki(Williams vd., 2018). Herhangi bir siber saldırının gerçekleşmeden önlenmesi bir siber güvenlik personeli için imkansızdır. Tepki aşaması ise saldırının gerçekleştiği ve hasarın meydana geldiği anlamına gelmektedir. Bu aşama hasar kontrolü ve kaybı en aza indirme üzerine odaklanır. Bu nedenle, çoğu güvenlik mekanizması, ülke incelemelerinde de belirtildiği üzere tehdit tespitine odaklanmaktadır. Örneğin, oltalama sayfasının oluşturulmasının engellenmesi mümkün olmamakla birlikte belirli bir alan adının şüpheli bulunması ve hızlıca tespit edilip engellenmesi mümkündür.

Ülke incelemeleri göstermektedir ki siber tehditlerin tespitinde iki farklı yaklaşımdan yararlanabilir: kural temelli ve anomali temelli. İlki, gelecek tehditlerin geçmişte olanlarla aynı kalıp ve örüntüleri sergileyeceği varsayılarak belirli bir tehdide karşılık gelen kalıp ve örüntülerin önceden tanımlanması ile çalışmaktadır. İkincisi, normal durum kavramının tanımlanmasını gerektirmektedir ve bu normallikten sapma gösteren olayları tespit etmeyi amaçlar. Bu sapmaların güvenlik olaylarına karşılık geldiği varsayılmaktadır. Bu iki tespit yaklaşımı genellikle birlikte kullanılmakta ve birbirini tamamlamaktadır. Kural temelli yaklaşımlar yüksek doğrulukla çalışmaktadır ancak sadece önceden maruz kalınan tehditleri tespit edebilir. Anomali temelli yaklaşımlar daha fazla yanlış sonuç üretme eğilimindedir ancak yeni saldırılara karşı daha yüksek başarı şansına sahiptir.

Makine öğrenmesinin ortaya çıkmasından önce, tespit mekanizmaları hem kural hem de anomali temelli olanlar için gerekli tüm öğelerin manuel olarak tanımlanmasını gerektirmekteydi. Bu durum, zaman alıcı ve hata yapmaya müsait bir görev olmanın yanı sıra, hızla artan tehdit vektörlerine karşı yetersiz kalmaktaydı. Ülke incelemelerimiz ve yapılan çalışmalar(Buczak & Guven, 2016) göstermektedir ki veri analitiği tekniklerinin ilerlemesiyle birlikte, makine öğrenmesinden yararlanılmaya başlanmıştır. Makine öğrenmesi, kullandığı modeller nedeniyle normalde insanlar tarafından fark edilemeyen detayları ve göstergeleri bularak daha iyi sonuçlar ortaya koymaktadır.

Siber tehdit tespiti için makine öğrenmesi uygulamalarının ayırt edici özelliği, denetimli veya denetimsiz makine öğrenmesi yöntemlerinin kullanılıp kullanılamayacağıdır. Önceden de belirtildiği gibi denetimli yöntemler, bütüncül tespit sistemleri olarak kullanılabilir ancak insan gözetimi ile oluşturulan etiketli verilere ihtiyaç duyarlar. Denetimsiz yöntemler ise insan müdahalesi olmadan çalışabilir ancak yalnızca üstünde çalıştıkları görevi yerine getirebilir, bütüncül sonuç sunamazlar. Denetimli öğrenme için analiz edilecek veri türüne bağlı olarak etiketlerin elde edilmesinde zorluk değişebilir. Örneğin, herhangi bir siber güvenlik uzmanı, yasal bir web sayfasını oltalama sayfasından ayırt edebilir ancak zararsız veya kötü niyetli ağ trafiğini ayırt etmek ve etiketlemek daha çok uzmanlık gerektirmektedir. Bu durumlarda denetimsiz eğitim daha avantajlı olabilmektedir.

Yapılan ülke incelemelerinde denetimli öğrenme ile kural temelli yaklaşımlara benzer çözümler üretilebilirken denetimsiz öğrenme ile sistemin normali tespit edilip anomali temelli çözümler üretilebildiği gözlemlenmiştir. Yapılan incelemeler tespit konusunda yapay zeka destekli üç farklı alan olduğunu göstermiştir; ağ saldırıları tespiti, zararlı yazılım tespiti ve oltalama tespiti.

3.2.1.1. Ağ saldırılarının tespiti

Kurumların siber güvenlik konusunda önem verdiği alanlardan biri saldırı tespitidir ve Saldırı Tespit Sistemleri (*Intrusion Detection Systems*-IDS) aracılığıyla

gerçekleştirilir. IDS'ler, genel olarak iki kategoride toplanırlar: Ağ Saldırı Tespit Sistemleri (*Network Intrusion Detection Systems*-NIDS), ağ seviyesindeki aktiviteleri analiz eder ve Ana Bilgisayar (Sunucu) Saldırı Tespit Sistemleri (*Host Intrusion Detection Systems*-HIDS), bireysel sunucu seviyesindeki aktiviteleri analiz etmektedir.

NIDS, ağ ortamında herhangi bir yerde konuşlandırılarak bulut, IoT, uç nokta cihazları ve endüstriyel kontrol sistemleri gibi çeşitli hedeflere yönelik tehditleri tespit etmek için makine öğrenmesinden yararlanabilir(Kshetri, 2021). Makine öğrenmesi ile desteklenen yararlanan bir NIDS, paket yakalama[26] (*packet-captures*-PCAP), ağ akışları (*NetFlows*[27]), Basit Ağ Yönetim Protokolü[28] (*Simple Network Management Protocol*-SNMP) veya DNS kayıtları gibi farklı türdeki verileri analiz edebilir. Özellikle modern ağların giderek büyümesiyle, NetFlow analizleri geleneksel PCAP'a göre birçok avantajı nedeniyle; gizlilik, depolama için daha az alan gereksinimi ve daha hızlı işleme süreleri, tercih edilmektedir(Yehezkel vd., 2021). Tüm ağ için etiketli veri elde etmek zor olduğu için denetimsiz öğrenme ile yetiştirilen Makine Öğrenmesi modelleri NIDS konusunda daha çok ön plana çıkmaktadırlar. Araştırmalar göstermektedir ki geleneksel yöntemlerle çalışan NIDS'lere göre makine öğrenmesi kullanan çözümler 4 kata kadar daha etkili olmaktadır(Apruzzese vd., 2017). NIDS'lerde denetimli makine öğrenmesi kullanmak da mümkündür. Bu yöntem kullanıldığında yüksek kalitede ve büyük miktarlarda etiketlenmiş veri gerekmektedir. Bu veriler ile eğitimden sonra çok iyi sonuçlar alınan sistemler üretilmesi mümkün olmaktadır. Bu sistemlerde hatalı doğru bulma oranı da düşüktür (Bilge vd., 2014).

---

[26]Bir bilgisayar ağındaki veri trafiğini yakalamak ve kaydetmek için kullanılan formatlardan biridir. PCAP dosyaları, ağdaki veri paketlerinin ham hallerini içeren ve bu paketleri daha sonra analiz etmek için kullanılabilen dosyalardır.

[27]Cisco tarafından geliştirilmiş bir ağ protokolüdür ve ağ trafiği hakkında detaylı bilgi toplamak ve analiz etmek için kullanılır. Ağ üzerinde hangi verilerin aktarıldığını, hangi kaynak ve hedef adreslerinin kullanıldığını ve trafiğin hangi uygulamalardan geldiği gibi bilgiler ihtiva etmektedir.

[28]Ağ yönetimi için kullanılan bir protokoldür. Ağ cihazlarını izlemek ve yönetmek için kullanılır. Ağ yöneticilerine ağ cihazlarıyla ilgili bilgileri toplama ve cihazları uzaktan kontrol etme imkanı sağlar.

3.2.1.2. Kötü Amaçlı Yazılım Tespiti

Kötü amaçlı yazılımlarla mücadele, siber güvenliğin en temel görevlerinden biridir. Kötü amaçlı yazılım belirli bir cihazı etkilediğinden, yazılımın tespiti ana bilgisayardaki veriler analiz edilerek yani HIDS'ler kullanılarak gerçekleştirilir. Antivirüs programları HIDS'in bir alt kümesi olarak kabul edilebilir. Kötü amaçlı yazılımlar, neredeyse her zaman belirli bir işletim sistemi[29] (*Operating Systems*-OS) için uyarlanmıştır. Windows işletim sisteminin popülaritesi nedeniyle yirmi yılı aşkın süredir en yaygın kötü amaçlı yazılımların hedefi olmuştur. Ancak günümüzde saldırganlar dikkatlerini mobil cihazlara yöneltmektedirler.

Kötü amaçlı yazılım tespiti, statik ve dinamik olmak üzere iki türde incelenebilir. Statik analiz, herhangi bir kod çalıştırmadan, sadece belirli bir dosyayı analiz ederek kötü amaçlı yazılımı tespit etmeyi amaçlamaktadır. Dinamik analiz ise bir yazılımın kullanılması sırasında ve genellikle kontrollü bir ortamda faaliyetler izlenerek gerçekleştirilen analiz ile kötü amaçlı yazılımı tespit etmeyi hedeflemektedir. Hem statik hem de dinamik analizler makine öğrenmesinden faydalanabilir.

Statik Analiz, daha basit olan yöntemdir. Özellikle bilinen kötü amaçlı yazılımlara karşı etkilidir ve birçok şekilde makine öğrenmesi ile geliştirilebilir. Örneğin, kümeleme, benzer kötü amaçlı yazılım özelliklerini tanımlamak için kullanılmaktadır(Hu vd., 2013). Makine öğrenmesi başarılı bir biçimde statik analiz metotları için kullanılsa da bu başarılara rağmen, tüm statik kötü amaçlı yazılım tespit yaklaşımları bilgisayar korsanları tarafından atlatılmaya açıktır. Kötü amaçlı yazılımın yürütülebilir dosyasını değiştirerek, altta yatan zararlı mantığı değiştirmeden statik

[29]Bilgisayar donanımı ile kullanıcılar arasında bir köprü görevi gören yazılımlardır. İşletim sistemi, donanım kaynaklarını yönetir, yazılım uygulamalarının çalışmasını sağlar ve kullanıcıların bilgisayarları etkili bir şekilde kullanmalarına yardımcı olur. İşletim sistemleri, masaüstü bilgisayarlar, sunucular, mobil cihazlar ve gömülü sistemler gibi her bilgisayarda kullanılmaktadır.

analizi atlatmak mümkündür. Gelişmiş kötü amaçlı yazılım varyantları (polimorfik[30] veya metamorfik[31]), yürütülebilir dosyalarını otomatik olarak değiştirerek statik tespit yaklaşımlarından kaçınabilirler.

Dinamik analiz yaklaşımlarının makine öğrenmesi teknikleri ile birleştirilmesi, polimorfik kötü amaçlı yazılımlara karşı etkili olmaktadır. Birçok makine öğrenmesi çözümü, kümeleme yönteminden yararlanır. Benzer davranış sergileyen kötü amaçlı yazılımları gruplandırmak, sadece daha önce görülmemiş kümelerin tespit edilmesine olanak tanımaktadır(Rieck vd., 2008). Dinamik analiz uygulanırken yapay zeka kullanılması ve buna ek olarak doğal dil işleme çalışmalarından da faydalanılması yöntemiyle de verimli sistemler oluşturulabilmektedir(Amer & Zelinka, 2020). Son olarak, statik analizle dinamik analizleri makine öğrenmesi aracılığıyla birleştirmek mümkündür. Bu sayede, denetimsiz öğrenme ile denetimli öğrenmeyi birleştirip yeni kötü amaçlı yazılımları tespit etmeyi sağlayan sistemler geliştirilerek, bu sistemlerin yüzde doksanın üzerinde tespit oranlarına ulaşması sağlanabilir(Chakraborty vd., 2020).

3.2.1.3. Oltalama Tespiti

Kimlik avı, hedef bir ağa sızmak için kullanılan en yaygın saldırı vektörlerinden biridir ve siber güvenlik ortamında hala yaygın bir tehdit olmaya devam etmektedir(Kettani & Wainwright, 2019). Kimlik avı girişimlerinin erken tespiti, kurumlar için büyük önem taşımaktadır ve makine öğrenmesinden büyük ölçüde fayda sağlayabilirler. Özellikle, oltalama girişimlerine karşı makine öğrenmesinin iki farklı uygulaması

[30]Polimorfik kötü amaçlı yazılımlar, her yayılışlarında veya çoğaldıklarında kendilerini farklı bir şekilde şifreleyerek veya kodlarını değiştirerek tespit edilmelerini zorlaştıran zararlı yazılım türleridir. Bu tür kötü amaçlı yazılımlar, şifreleme anahtarlarını ve şifre çözme rutinlerini değiştirerek her yeni kopyasında farklı bir görünüme sahip olur.

[31]Metamorfik kötü amaçlı yazılımlar, kendilerini yeniden yazıp değiştirerek her çoğaltıldıklarında veya yayıldıklarında tamamen farklı bir kod yapısına sahip olacak şekilde tasarlanmış zararlı yazılım türleridir. Bu zararlı yazılımlar, kodlarını anlamlı bir şekilde yeniden düzenleyebilir, talimatları yeniden sıralayabilir ve hatta farklı programlama teknikleri kullanarak kendilerini yeniden yazabilirler.

bulunmaktadır; yasal bir web sitesini taklit edecek şekilde kamufle edilmiş web sayfalarını tanımlamayı hedefleyen oltalama sitelerinin tespiti veya tehlikeli bir web sitesine yönlendiren ya da hassas bilgileri içeren bir yanıt gönderilmesini teşvik eden oltalama e-postalarının tespiti. Bu iki yaklaşım arasındaki temel fark, analiz edilen veri türüdür. Web siteleri için, sayfanın Evrensel Kaynak Konumlayıcı[32] (*Universal Resource Locator*-URL) adresi, Hiper Metin İşaretleme Dili[33] (*Hyper Text Media Language*-HTML) kodu veya hatta görsel temsili kullanılırken(Tian vd., 2018), e-postalar için, tipik olarak e-postanın metni, başlığı veya ekleri analiz edilir(Alhogail & Alsabih, 2021).

Oltalama web siteleri çoğunlukla kara listeler aracılığıyla engellenmekle birlikte bu listeler hızla güncelliklerini kaybetmektedir. Çünkü saldırganlar oltalama tuzaklarını sık sık bir siteden diğerine taşımaktadır ve bu taşımayı yapan oltalama web sitelerinin yüzde doksanından fazlası popüler kara listeler tarafından tespit edilememektedir(Tian vd., 2018). Makine öğrenmesi, manuel ve statik kara listelemeye karşı ülkemizde de kullanılan bir alternatiftir ve çoğu modern web tarayıcı da bu alternatiften yararlanmaktadır (Liang vd., 2016a). Kötü amaçlı yazılım veya ağ saldırı tespitiyle karşılaştırıldığında, kimlik avı web sitelerine karşı denetimsiz makine öğrenmesi daha az kullanılır. Yapılan bir çalışmada görseller, HTML ve URL kullanılarak yaklaşık 1.000 tane sürekli alan adı değiştiren oltalama web sitesi için, manuel kara listeleme yalnızca yüzde dokuzunu tespit ederken, makine öğrenmesi oltalama girişimlerinin yüzde yetmişini tespit etmiştir(Tian vd., 2018).

Kimlik Avı E-postası Tespiti makine öğrenmesinin siber güvenlik için ilk uygulamalarından biri olarak karşımıza çıkmaktadır. Bu tespit sistemi, istenmeyen e-postaların (*spam*) tespitinde kullanılmıştır. Doğal dil işleme alanındaki gelişmeler,

---

[32]İnternet üzerinde bir kaynağın adresini belirten standart bir tanımlayıcıdır. Bir web sitesine, dosyaya, resme veya diğer çevrimiçi kaynaklara erişmek için kullanılır. URL, genellikle üç ana bileşenden oluşur: protokol (örneğin, http veya https), alan adı (örneğin, www.example.com) ve kaynak yolu (örneğin, /index.html).

[33]Web sayfalarını oluşturmak ve yapılandırmak için kullanılan standart bir işaretleme dilidir.

makine öğrenmesi tarafından e-postanın içeriğini analiz etmek ve kötü niyetli amacı tespit etmek için kullanılabilmektedir (Biggio & Roli, 2018).

Denetimsiz makine öğrenmesi web sitesi örneğinde olduğu gibi e-posta alanında da sık kullanılmamakla birlikte, yüzde doksan beşin üzerinde tespit oranına ulaşan denetimsiz makine öğrenmesi algoritmaları geliştirilmiştir(Diale vd., 2019). Ancak, oltalama web sitesi tespitinde olduğu gibi e-postalar için doğru etiketlerin elde edilmesi nispeten daha kolay olmakla e-posta sağlayıcıları tarafından denetimli makine öğrenmesi kullanan otomatik filtrelerin geliştirmesini de kolaylaştırmaktadır (Karim vd., 2019).

Son olarak, oltalama ve spam ile mücadele son zamanlarda çevrimiçi sosyal ağlara da taşınmıştır. Bu ortam, doğal dil işlemeden faydalanıldığı için e-postalardaki kimlik avı tespiti ile büyük benzerlik göstermektedir. Örneğin, derin öğrenme kullanılarak kötü niyetli tweetler tespit edilmiş ve yüzde doksan beşe yakın tespit oranı ile yüzde beş hatalı pozitif oranı elde edilmiştir(Wu vd., 2017). Benzer şekilde, başka bir çalışmada özellikle kurbanları oltalama web sitelerine yönlendiren tweetlere odaklanılmış ve yüzde doksan beşin üzerinde tespit oranı ve yaklaşık yüzde doksan doğruluk elde edilmiştir(Lancaster vd., 2018).

3.2.2. Diğer Kullanım Alanları

Makine öğrenmesi, siber güvenlikte tehdit tespitinin yanı sıra birçok ek rol de üstlenebilmektedir. Modern siber dünya, önceki bölümde tanıtılan makine öğrenmesi modellerinin ürettiği veriler de dahil olmak üzere, çeşitlik kaynaklardan gelen büyük miktarda veriyi sürekli olarak üretmektedir. Bu verilerin makine öğrenmesi ile analiz edilmesi, dijital sistemlerin güvenliğini daha da artıracak bilgiler sağlayabilmektedir. Bu kadar büyük ve etiketlenmemiş bir veri havuzu olduğu zaman bu bölümde bahsedilen makine öğrenmesi uygulamalarının çoğunun ortak özelliği, denetimsiz makine öğrenmesi kategorisinde bulunmalarıdır. Bunun nedeni de uygulamaların geniş kapsamlı ve insan yönlendirmeli etiketleme prosedürlerini gerektirmemeleridir.

3.2.2.1. Uyarı Yönetimi

Yapay zeka olsun ya da olmasın kusursuz bir tespit sistemi geliştirmek mümkün değildir. Bu nedenle, yanlış tahminlere dayalı eylemlerin otomatik olarak gerçekleştirilmesini önlemek için tespit sistemlerinin çıktıları genellikle uyarılar şeklinde olmaktadır. Bu uyarılara bağlı olarak daha uygun bir müdahale gerçekleştirilebilir. Ancak, modern sistemler her saat binlerce uyarı üretmektedir. Bu da el ile ayıklamayı imkansız hale getirmektedir. Bu sorunu çözmek için, makine öğrenmesi modellerini uyarıları filtrelemek, önceliklendirmek veya gerektiğinde daha genel bir olayla birleştirmek için kullanılabilir.

- Filtreleme: Her zaman bir uyarı zararlı bir olayı göstermemekte ve bir kısmı yanlış alarm olarak karşımıza çıkmaktadır. Siber olay olmayan farklı uyarılar tarafından teyakkuza geçirilmek, siber güvenlik uzmanlarının zamanının boşa harcanması anlamına gelmektedir. Bu nedenle, makine öğrenmesi aynı gerekçe ile oluşturulan uyarıları tespit edip gruplama gibi yöntemlerle, gereksiz uyarıları filtrelemekte yardımcı olmaktadır. Örnek vermek gerekirse üretilen yanlış alarmlar için özel olarak tasarlanmış makine öğrenmesi modeli gerçek botnet[34] trafiği üzerinde test edildiğinde, yanlış alarmların ayıklanmasına harcanan zamanı yüzde yetmiş beş oranında azaltmış ve makine öğrenmesi olmayan çözümlere göre dışı yüzde kırk beş daha etkili olmuştur(Su vd., 2019).
- Önceliklendirme: Siber güvenlik uzmanları her gün binlerce uyarı ile karşı karşıya kalmaktadır. En kritik uyarıların belirlenmesi için önceliklendirme teknikleri uygulanmaktadır. Makine öğrenmesi, denetimli öğrenme modelleri ile en ilgili sıralama kriterlerini otomatik olarak öğrenebilmektedir. Makine öğrenmesi destekli önceliklendirme modelleri en önemli uyarıları yüzde doksan beş oranında doğrulukla en üst sırada önceliklendirmeyi sağlayabilmektedir(Vidovic vd., 2021).

---

[34]"*Robot*" ve "*Network*" kelimelerinin birleşiminden oluşan bir terimdir. Bir grup bilgisayarın (*bot*) kötü niyetli bir saldırgan tarafından uzaktan kontrol edilmesiyle oluşturulan bir ağdır. Bu bilgisayarlar genellikle kötü amaçlı yazılımlar aracılığıyla sahiplerinin bilgisi dışında botnet ağına katılmaktadır.

- Birleştirme: Büyük miktarda uyarıyı yönetmenin verimli bir yolu da benzer uyarıları bir araya getirmek ve ardından çeşitli siber güvenlik olayları ile aralarındaki ilişkileri belirlemek için bu gruplar arasındaki korelasyonları bulmaktır. Kötü niyetli aktörlerin saldırmayı tercih ettikleri protokoller ve ağ portlarını belirlemek için kümeleme geçekleştiren yapay zeka modelleri kullanılmıştır(Okutan & Yang, 2019). Sonuç olarak günümüz saldırılarının giderek daha fazla Uzaktan Masaüstü Protokolü[35]'ne (*Remote Desktop Protocol-RDP*) dayandığı ve bunun, yanal hareket faaliyetlerine olanak tanıdığını vurgulamaktadır.

Ülke incelemelerinde bahsedilen bazı ürünlerde olduğu gibi yukarıdaki tekniklerin tümü birleştirilebilir. Bu bağlamda, derin öğrenmeyi kullanarak uyarıları birleştiren ve önceliklendiren uyarı yönetim çözümleri test edilmiş ve gerçek siber güvenlik uzmanları tarafından kullanılabilir bulunmuştur(McElwee vd., 2017).

3.2.2.2. Ham Veri Analizi

Siber güvenlik alanı, her biri farklı nitelikte ham veri (loglar, raporlar, uyarılar) üreten sistemlerden toplanan veriyi anlamlandırmak ve incelemek zorundadır. Bu verilerin sağladığı faydanın en üst düzeye çıkarılması için makine öğrenmesinin yeteneklerinden yararlanılabilmektedir. Bu bağlamda makine öğrenmesinin iki uygulama alanı üzerinde durulabilir; günlük veri analizleri yoluyla operasyonel kararların desteklenmesi ve verilerin etiketlenmesi ile ilgili çalışmaları optimize ederek denetimli makine öğrenmesi modellerinin eğitilmesi için kullanılması.

- Operasyonel kararların desteklenmesi; bilgi sistemlerinde günlük verilerinin bolluğu ve bu verilerin eğitimde kullanılabilme potansiyeli operasyonel güvenlik bağlamında makine öğrenmesini ön plana çıkartmaktadır. Örnek olarak, günlük verilerinden (*proxy*[36], Dinamik Ana Bilgisayar Kontrol

[35]Microsoft tarafından geliştirilen bir ağ iletişim protokolüdür. RDP, kullanıcıların başka bir bilgisayara veya sunucuya uzaktan bağlanarak o bilgisayarı sanki önündeymiş gibi kullanmalarını sağlamaktadır.

[36]Proxy server, istemci bilgisayarlar ile hedef sunucular arasında aracılık yapan bir sunucudur.

Protokolü[37] (*Dynamic Host Control Protocol*-DHCP) veya Sanal Özel Ağ[38] (*Virtual Private Network*-VPN) sunucuları tarafından üretilen) bilgi çıkarımına odaklanan ilk denetimsiz makine öğrenmesi sistemlerinden biri Beehive'dır(T. F. Yen vd., 2013). Çalışmadaki amaç, logların anomali tespitinde kullanılarak müdahale gerektiren siber olayların bu olaylara karşılık gelen log kalıplarıyla ilişkilendirilmesiydi. Beehive, EMC Corporation[39]'ın iki haftalık günlük verilerini değerlendirmiştir. Neredeyse 800 olayı tespit etmiştir, bunların yüzde altmış beşi gerçek güvenlik olaylarıyla ilişkili çıkmıştır. Karşılaştırıldığında, makine öğrenmesi olmayan yöntemler çok daha kötü performans göstermişler ve sadece 8 doğru olay tespit edebilmişlerdir. Derin öğrenmenin ortaya çıkışıyla geliştirilen başka bir örnek DeepLog'dur(Du vd., 2017), Beehive ile benzer bir amaçla günlük verilerini (Hadoop[40] veya OpenStack[41] günlükleri) analiz etmektedir. DeepLog, mevcut verilerin sadece yüzde biri ile eğitim aldıktan sonra laboratuvar ortamında neredeyse yüzde yüz tespit oranıyla etkileyici sonuçlar elde etmektedir.

- Etiketleme optimizasyonu; önceki bölümlerde de incelediğimiz üzere birçok tehdit tespit tekniği denetimli makine öğrenmesi yöntemleri kullandığı için büyük miktarda etiketli veri gerektirmektedir. Buna karşılık, etiketlenmemiş veriler siber güvenlikte çok yaygındır ve birçok çalışma, küçük etiketli veri setlerinin getirisini artırmak ve böylece denetimli makine öğrenmesi yöntemlerinin uygulanmasını sağlamak için yarı denetimli öğrenme yöntemleri önermektedir(Apruzzese vd., 2022). Örneğin, botnet tespiti için geliştirilmiş bir makine öğrenmesi modeli(Y. Zhang vd., 2021) sadece 2.400 etiketli

[37]Ağdaki cihazlara dinamik olarak IP adresleri ve diğer ağ yapılandırma bilgilerini atayan bir ağ yönetim protokolüdür.

[38]İnternet üzerinden güvenli ve şifreli bir bağlantı oluşturarak, kullanıcıların ve cihazların özel bir ağdaymış gibi çalışmasını sağlayan bir protokoldür.

[39]EMC Corporation, veri depolama, bilgi güvenliği, sanallaştırma, analiz, bulut bilgi işlem ve diğer bilgi teknolojisi hizmetleri ve çözümleri sunan bir Amerikan çok uluslu şirkettir.

[40]Büyük veri kümelerinin depolanması ve işlenmesi için açık kaynaklı bir yazılım çerçevesidir.

[41]Bulut bilişim altyapısı oluşturmak için kullanılan açık kaynaklı bir yazılım platformudur.

örnekle 0.83 F1-skora[42] ulaşırken, başka bir makine öğrenmesi modeli(Apruzzese vd., 2020) aynı ağ senaryosunda 0.95 F1-skoruna ulaşmak için milyonlarca etiketli örnek kullanmıştır. Bu bilgilere paralel olarak yürütülen bazı çalışmalarda ise; küçük bir etiketli veri seti üzerinde eğitilmiş bir makine öğrenmesi modeli kullanılarak büyük bir etiketlenmemiş veri setinde hangi örneklerin etiketlenmesi gerektiğinin önerilmesi ve böylece öğrenme oranının iyileştirilmesi de incelenmiştir (Pendlebury vd., 2018; X. Zhang vd., 2020).

3.2.2.3. Risk Maruziyet Değerlendirmesi

Herhangi bir siber saldırının tamamen önlenmesi ulaşılamaz bir hedef olsa da bir sistemin zayıf noktalarına odaklanarak ve en olası tehditleri önceden tahmin ederek önemli ölçüde güçlendirilmesi mümkündür. Bu bağlamda, makine öğrenmesi çeşitli görevlerde (sızma testi veya tehlike göstergelerinin tahmin edilmesi) yardımcı olabilmektedir.

Güvenlik sistemlerine otomatik olarak saldırarak sızma testi gerçekleştiren bir makine öğrenmesi modeli güvenlik açığı değerlendirmesi için faydalı bir varlık olabilecektir. Geleneksel NIDS'lere karşı sentetik saldırılar oluşturmak için kullanılan bir makine öğrenmesi modeli, manuel inceleme süresinin yarısında aynı miktarda güvenlik açığı bulmuş ve rastgele saldırı prosedürlerine göre yüzde doksan hızlanma sağlamıştır(Ghanem & Chen, 2018). Makine öğrenmesi tabanlı bir botnet tespit sistemini otomatik olarak atlatmak ve ardından edinilen tecrübeyle güçlendirmek için

[42]F1-score, makine öğrenmesi ve istatistikte kullanılan bir performans ölçütüdür, kesinlik (*precision*) ve duyarlılık (*recall*) gibi metriklerin bir kombinasyonudur. Kesinlik, doğru pozitif tahminlerin, toplam pozitif tahminlere oranıdır $precision = \frac{TP}{TP+FP}$. Duyarlılık ise doğru pozitif tahminlerin, toplam gerçek pozitiflere oranıdır $recall = \frac{TP}{TP+FN}$.

$$F_1 Score = 2 \times \left(\frac{precision \times recall}{precision + recall}\right)$$

derin öğrenme yaklaşımı ile çalışan bir makine öğrenmesi modeli de bulunmaktadır(Apruzzese vd., 2020). Benzer şekilde, makine öğrenmesi ile oluşturulmuş SQL enjeksiyon[43] saldırılarına karşı veri tabanlarının güvenlik açıklarını sınayan makine öğrenmesi projeleri de geliştirilmiştir(Uwagbole vd., 2017). Tüm bu çalışmaları tek bir modelden gerçekleştirmek için makine öğrenmesi destekli platform önerileri de bulunmaktadır(Chaudhary vd., 2020). 2019 yılında yapılan bir araştırmaya göre, sızma testi için makine öğrenmesinin potansiyeli hala geliştirmeye açık bir alandır(McKinnel vd., 2019).

Tehlike göstergelerinin tahmin edilmesi için makine öğrenmesinden faydalanılabilir. Bir sistemdeki bilgisayarların hangi olasılıkla tehlikeye maruz kaldığını tahmin eden makine öğrenmesi modelleri üzerinde çalışmalar yapılmaktadır. Yapılan bir çalışmada kurumsal bir ağ ortamı incelenerek, makine öğrenmesi her bir ana bilgisayarın ve tüm ağın davranışlarını analiz etmek için kullanılmıştır(T.-F. Yen vd., 2014). Bu analizler için kullanılan veriler uç nokta koruma cihazları ve her ana bilgisayarın belirli kullanıcısına ait kişisel bilgileri de içermektedir. Bulgular, iş web sitelerini ziyaret etmenin yüzde otuzluk bir oranla, tehlikeye maruz kalmış ana bilgisayarların en yaygın göstergesi olduğunu ve ikinci sırada ise yüzde on beşle seyahat web sitelerinin yer aldığını ortaya koymuştur. Makine öğrenmesini bal-küpü[44] (*honeypots*) ile birleştirmesi de üzerinde çalışılan uygulamalar arasındadır(Martínez Garre vd., 2021). Bu strateji, *botnet* kötü amaçlı yazılımı tarafından enfekte olma olasılığı yüksek ana bilgisayarları belirlemek için kullanılmaktadır. Ayrıca, Facebook da farklı kaynaklardaki verileri kullanarak sahte hesapları belirlemek için makine öğrenmesinden yararlanmaktadır ve bu çalışma sonucunda kullanıcıların yüzde otuz oranında daha az sahte hesap ile karşılaşacağı gösterilmiştir(Xu vd., 2021)

---

[43]Bir siber saldırı türüdür, web uygulamalarındaki SQL sorgularını manipüle ederek, saldırganın veri tabanına izinsiz erişim sağlamasına veya veri üzerinde yetkisiz işlemler yapmasına olanak tanır.

[44]Siber güvenlikte, saldırganları tespit etmek, analiz etmek ve yönlendirmek için kullanılan tuzak sistemlerdir. Gerçek sistemleri taklit eden bu güvenlik araçları, saldırganların dikkatini çekerek onların yöntemlerini ve hedeflerini ortaya çıkarır. Böylece, güvenlik uzmanları saldırıları inceleyebilir ve gerçek sistemlerin güvenliğini artırabilirler.

3.2.2.4. Tehdit İstihbaratı

Tehdit istihbaratının ana görevi, yeni saldırıları öngörmek için bilgi toplamaktır. Bu çoğu siber savunma yönteminin aksine proaktif bir süreçtir ve savunmaları son saldırılara karşı güncel tutmak için güçlü bir araçtır. Tehdit istihbaratı için makine öğrenmesi uygulamaları hem iç hem de dış veri kaynaklarından yararlanabilmektedir. Makine öğrenmesi kullanarak gelecekteki saldırı stratejilerini öngörmek için iç verilerden yararlanılabilir. Makine öğrenmesi modeli geçmiş siber saldırılara karşılık gelen uyarılar oluşturmak ve ardından bu uyarıları bir saldırganın davranışını incelemek için kullanmak amacıyla eğitilebilmektedir(Sweet vd., 2020). *SAGE* isimli bir proje, makine öğrenmesi kullanarak üç yüz binden fazla bireysel uyarıyı kullanarak, bunları yaklaşık yüz saldırı modelinden biriyle eşleştirecek şekilde eğitilmiştir(Nadeem vd., 2021).

Bir diğer kullanım alanı ise, yürütülebilir dosyaları parçalarına ayırmak için derin öğrenmeyi kullanmak ve bu sayede gelecekteki kötü amaçlı yazılımlarda yeniden ortaya çıkabilecek ve potansiyel olarak kötü niyetli olduğu tespit edilen bazı kalıpların tanımlanmasını sağlamaktır. EKLAVYA isimli bir proje bu görevde yüzde seksen gibi bir doğruluk oranına ulaşmaktadır(Chua vd., 2017). Bir başka projede ise gelecekteki kötü amaçlı yazılımların bir kurumu nasıl etkileyebileceğini öngörmek için daha önceden karşılaşılan kötü amaçlı yazılım bilgileri bir makine öğrenmesi modeli eğitmek için kullanılmıştır. Bu model, makine öğrenmesi olmayanlara göre 4 kat daha fazla doğru tahmin sağlamıştır(Kang vd., 2016).

Makine öğrenmesi ayrıca açık kaynak istihbaratın değerlendirilmesi için de kullanılabilir. Twitter'da bahsedilen güvenlik olaylarına odaklanan bir çalışmada makine öğrenmesi modelleri 2016 yılında meydana gelen birçok kötü niyetli faaliyeti tespit etmiştir. Örneğin; Mirai botnet[45] (Ekim 2016) veya AdultFriendFinder veri

[45]İnternet bağlantılı cihazları hedef alan ve onları birer bot haline getirerek büyük ölçekli DDoS saldırıları düzenlemek için kullanılan kötü amaçlı bir yazılımdır.

ihlali[46] (Kasım 2016), gibi olaylar bu şekilde tespit edilmiştir (Sapienza vd., 2017). Benzer şekilde derin öğrenme yöntemi ile fidye yazılımı saldırılarının gelişimini incelemek için ilgili tweetler analiz edilmiştir(R. vd., 2019). Ortak Güvenlik Açığı Skoru[47] (*Common Vulnerability Score*-CVS) gibi güvenlik geri bildirimlerinden gelen bilgileri kullanan projeler de bulunmaktadır. Makine öğrenmesi kullanarak CVS'leri geleneksel siber güvenlik yöntemlerinden yaklaşık bir hafta önce tahmin eden modeller geliştirilmiştir(Chen vd., 2019). Makine öğrenmesi ile CVS'nin tahmini, *darkweb*[48] verileri üzerinden de yapılabilir(Almukaynizi vd., 2017). Bu çalışmada yer altı forumları taramak ve anlamlı bilgileri güvenlik açığı açıklamalarıyla ilişkilendirmek için makine öğrenmesi kullanılmaktadır. Sonuçları üçüncü taraf kaynaklar ile doğrulayan çalışmada, önerilen makine öğrenmesi yöntemi, istismar edilebilir güvenlik açıklarının yaklaşık yüzde kırkını tespit edebilmiştir, geleneksel yöntemler ise yaklaşık yüzde onunu tespit etmiştir. Kötü niyetli yazılımlar ve yöntemlerin fiyatlarının tespit edilmesi için yer altı forumlarının siber suç pazarlarını ortaya çıkarmayı amaçlayan başka bir makine öğrenmesi modeli de geliştirilmiştir(Portnoff vd., 2017).

3.3. Makine Öğrenmesi ile ilgili sorunlar

Makine öğrenmesi, bağımsız ve aynı şekilde dağıtılmış rastgele değişkenler (*independent, identically distributed random variables*-iid) ilkesini takip eder(Dundar vd., 2007). Bu ilke, makine öğrenmesi modeli geliştirirken analiz edilen verilerin, modelin eğitimi tamamlandıktan sonra modelin analiz edeceği gelecek verilerle benzer olacağının varsayılmasıdır. Eğer iid varsayımı doğru olmazsa, oluşturulan makine

[46]Kasım 2016'da gerçekleşen büyük bir siber saldırıdır. Bu ihlal sonucunda yaklaşık 412 milyon kullanıcı hesabının bilgileri sızdırılmıştır. Sızdırılan veriler arasında kullanıcı adları, şifreler, e-posta adresleri ve diğer kişisel bilgiler yer almaktadır. Bu ihlal, AdultFriendFinder Network'e ait birkaç farklı sitenin veri tabanlarını etkilemiştir.

[47]Bilgisayarın güvenlik açıklarını değerlendirmek ve önceliklendirmek için kullanılan standart bir puanlama sistemidir.

[48]İnternet üzerinde doğrudan erişilebilir olmayan, normalde endekslenmeyen, erişmek için özel yazılım, yöntem ve konfigürasyonların gerektiği, internet ağı.

öğrenmesi modeli, geliştirme sırasında ölçülen performanstan farklı bir performans sergileyecektir. Bu ilke, siber güvenlikte makine öğrenmesinin kullanımını kısıtlayabilmektedir. Çünkü bu alanın bazı özellikleri ile çelişmektedir. Siber güvenlik tehditleri sürekli evrim geçirmektedir. Bu, bir makine öğrenmesi modelinin eğitim verileri ile gerçek dünya verileri arasında uyumsuzluk yaratabilmektedir. Yeni tehditler ve saldırı yöntemleri ortaya çıktıkça, modelin performansı düşebilir. Siber güvenlik tehditleri, saldırganların makine öğrenmesi modelini yanıltmak veya atlatmak için aktif olarak çalıştığı düşmanca bir ortamda gerçekleşmektedir. Saldırganlar, modelin zayıf noktalarını belirleyip bunlardan yararlanabilirler. Bu da modelin güvenilirliğini azaltacaktır. Siber güvenlik verileri genellikle yüksek derecede gizlidir ve bu verilerin gizliliğinin korunması gerekmektedir. İncelendiği üzere makine öğrenmesi modellerinin eğitimi için büyük miktarda veri gerekebilir ancak bu verilerin gizliliğini korumak da zor olabilir. Verilerin paylaşımı veya transferi sırasında gizlilik ihlalleri meydana gelebilmektedir. Bu özellikler, siber güvenlik alanında makine öğrenmesi modelinin uygulanmasını zorlaştırabilmektedir.

Günümüz sistemleri, her gün eklenen yeni cihazlar, hizmetler ve hatta kullanıcılar ile sürekli kendini yenilemektedir. Bu tür tüm değişimler, iid varsayımı ile uyuşmayan durumlar ortaya çıkartmaktadır. Eğitim verileri hızla güncelliğini yitirdiği için uzun vadede makine öğrenmesinin güvenilirliğinde azalma yaşanacaktır. Bu sorun her ne kadar makine öğrenmesinin herhangi bir uygulamasını etkileyebilse de bazı alanlar bu durumdan daha az etkilenmektedir. Örnek vermek gerekirse, yazı okuma gibi bir uygulamada harfler her zaman aynı gözükecek ve kelimeleri oluşturacaklardır. Bu nedenle aynı verilerle eğitilmiş bir makine öğrenmesi modelini uzun süre kullanmak mümkün olacaktır. ImageNet isimli ve görüntü işleme modellerini eğitmek için kullanılan bir veri seti 2011 yılından bu yana hala kullanılmaktadır (Ramanathan vd., 2021). Siber güvenlikte bu kadar uzun vadeli kullanım mümkün değildir. Siber güvenlik doğası gereği sürekli değişmekte ve kendini yenilemektedir.

Yeni bir güvenlik açığı keşfedildiği zaman, daha önce zararsız kabul edilen bazı olayların zararlı olarak ele alınması gerekmektedir. Ya da bir ağ yeni bir cihaz ya da protokol kullanılarak genişletildiğinde, ağın diğer taraflarından farklı davranışlar

sergileyerek birçok yanlış anomali yaratabilmektedir. Ayrıca, saldırganlar mevcut sistemler tarafından tespit edilmemiş sıfırıncı gün açıklıkları gibi yeni yöntem ve stratejiler geliştirmektedir. Araştırmalar, bu sorunun kaynağı ve makine öğrenmesi tespit modellerindeki performans düşüşleri üzerinde yoğunlaşmaktadır (Andresini vd., 2021; Jordaney vd., 2017). Bu soruna karşı bir çözüm ise, makine öğrenmesi modellerinin eğilimleri için kullanılan verilerin günceli yansıtan yeni verilerle sürekli desteklenmesidir.

Siber güvenlik alanı, tanımı gereği saldırgan tarafların olduğu bir alandır. Çoğu saldırı daha basit ve eski yöntemler kullansa da motivasyonu ve kaynağı yüksek saldırganlar sürekli olarak saldırı stratejilerini geliştirmekte ve değiştirmektedir. Sıfır gün saldırıları riskinin yanı sıra, makine öğrenmesinin kendisi de saldırılara maruz kalabilmektedir(Srndic & Laskov, 2014a). Bu tarz tehditler, makine öğrenmesi modelinin tahminlerini bozma amacıyla bazı verilerde küçük bozulmalar uygulanması ile gerçekleştirilmektedir. Algılanamayacak kadar küçük değişiklikler bile makine öğrenmesi modellerinin performansını etkileyebilmektedir. Örneğin, bazı ağ verilerine birkaç baytlık gereksiz veri eklenerek birçok makine öğrenmesi destekli botnet tespit sistemi atlatılabilmiştir(Apruzzese vd., 2019). Makine öğrenmesi destekli kötü amaçlı yazılım tespit sistemlerine karşı da benzer yöntemler kullanılmıştır(Pierazzi vd., 2019; Srndic & Laskov, 2014b). Google Chrome'un kimlik avı tespit sistemi gibi ticari ürünlerin üstündeki çalışmalarda da benzer durumlar gözlemlenmiştir(Liang vd., 2016b).

Siber güvenlik alanı doğası gereği veri gizliliğine önem verilmesi gereken bir alandır. Bu durum, makine öğrenmesinin eğitilmesi için gerekli olan verinin sağlanmasında engel teşkil edebilmektedir. Şifrelemenin artan kullanımı, bazı makine öğrenmesi sistemlerini tamamen kullanılamaz hale getirebilmektedir. HTTP trafiğini inceleyen bir makine öğrenmesi modeli, trafik HTTPS ile şifrelendiğinde çalışmaz hale gelecektir ve dünya genelinde şifreli iletişim artış göstermektedir. Bu sorun, kimlik avı e-postalarını tespit eden modeller gibi makine öğrenmesinin diğer kullanım alanlarını da etkileyebilmektedir. E-postalar şifrelenmişse, içeriklerinin modeller tarafından analiz edilmesi imkansız hale gelmektedir.

Makine öğrenmesinin karşılaşabileceği bir diğer sorunlu senaryo ise, gizli verilerin analiz edilmesini gerektiren durumlardır. Veri düzenlemelerindeki değişiklikler GDPR[49] gibi düzenlemeler uzun vadede güvenilir bir şekilde kullanılabilecek verileri tanımlamayı zorlaştırır. Zararlı yazılım enfeksiyon riskini tahmin etmek için diğer bilgilerin yanı sıra kullanıcı bilgilerini kullanan yaklaşımlar bulunmaktadır(T.-F. Yen vd., 2014). Bu tür yaklaşımlar, GDPR kapsamında kullanılacağı kurumun tüm kullanıcılarının açık onayının alınmasını gerektirmektedir. Ayrıca, gizlilik ve şifreleme, etiketleme prosedürlerini de zorlaştırmaktadır. Bir örneğin gerçeğe uygunluğunu doğrulamak için insan uzman tarafından görülmesi gerekmektedir. Son olarak, kurumlar verilerini kamuya açıklamak istemeyebilmektedir. Bu durum, makine öğrenmesi modellerinin değerlendirilmesi için kullanılabilecek kamuya açık veri setlerinin seyrekleşmesi ile sonuçlanabilmektedir(Sharafaldin vd., 2018).

[49]*General Data Protection Regulation*, Avrupa Birliği tarafından 25 Mayıs 2018'de yürürlüğe sokulan, bireylerin kişisel verilerini korumayı amaçlayan bir yasadır.

SONUÇ VE ÖNERİLER

SONUÇ

Makine öğrenmesi ve yapay zeka teknolojilerinin siber güvenlik alanında kullanımı, bu teknolojilerin ülkemizdeki durumu ve ülkemiz için sağlayabileceği faydalar incelenmiştir. Yapılan araştırmalar ve incelemeler sonucunda, makine öğrenmesi ve yapay zeka teknolojilerinin siber güvenlik tehditlerini tespit etme ve bu tehditlere karşı önlem alma noktasında oldukça etkili araçlar olduğu ortaya konulmuştur.

Makine öğrenmesi ve yapay zeka teknolojileri, büyük veri setlerini analiz ederek anormallikleri ve potansiyel tehditleri hızlı bir şekilde tespit edebilme kapasitesine sahiptir. Bu teknolojiler sayesinde, sürekli öğrenen ve güncellenen sistemler oluşturularak yeni ve gelişen tehditlere karşı proaktif bir savunma sağlanabilmektedir. Özellikle, ağ saldırılarının tespiti, kötü amaçlı yazılım tespiti, oltalama tespiti, uyarı yönetimi, ham veri analizi, risk maruziyet değerlendirmesi ve tehdit istihbaratı gibi siber güvenlik alanlarında yapay zeka ve makine öğrenmesi uygulamalarının büyük bir potansiyele sahip olduğu görülmüştür. Bu teknolojiler kullanılarak tespit edilen tehditlere karşı otomatik olarak tepki veren müdahale mekanizmaları geliştirilebilir ve böylece tehdit düzeyine göre farklı müdahale stratejileri uygulanarak insan etkisi en aza indirilebilir.

Tez kapsamında, Amerika Birleşik Devletleri, Birleşik Krallık, Fransa ve Türkiye'nin siber güvenlik uygulamaları incelenmiş ve bu ülkelerin makine öğrenmesi uygulamaları detaylı bir şekilde ele alınmıştır. Ülkelerin yapay zeka kullanımı ve ülkelerde geliştirilen ürünler ve yapılan araştırma geliştirme faaliyetleri göstermektedir ki önümüzdeki dönemlerde kamu kurumları ve kritik altyapılarda güvenliğin sağlanmasında yapay zekanın önemi giderek artacak ve kullanımı yaygınlaşacaktır.

Bu teknolojilerin KULLANIMI VE VERİMLİLİĞİ AÇISINDAN büyük miktarda kaliteli veri gerektiği anlaşılmaktadır. Toplanan verilerin güvenli bir şekilde

saklanması ve işlenmesi ile gizlilik ilkelerine uygun olarak kişisel verilerin korunması da sağlanmalıdır.

ÖNERİLER

Ülkemiz için makine öğrenmesinin siber güvenlik alanında etkili bir şekilde kullanılabilmesi hem milli güvenliğimiz hem de refah seviyemizin arttırılmasında faydalı olacaktır. Bu yolla ülkemizin siber saldırılardan kaynaklı zararları azaltılacak ve yapay zeka gibi sürekli önemi artan bir alanda ülkemizin güvenliğine, gelişimine ve farkındalığına katkı sağlanacaktır.

Eğitim ve Farkındalık

Makine öğrenmesi ve siber güvenlik konularında uzmanlaşmış profesyoneller yetiştirmek, ülkemizin bu alandaki başarısı için konjektürel temel bir gerekliliktir. Üniversiteler, bu alandaki eğitim programlarını genişleterek lisans ve lisansüstü düzeyde makine öğrenmesi ve siber güvenlik dersleri sunmaktadır. Fakat, bu dersler sadece teorik bilgi ile sınırlı kalmamalı, aynı zamanda pratik uygulamalar ve projelerle desteklenmelidir. Üniversitelerin yanı sıra her türlü kurumda da makine öğrenmesi ve siber güvenlik alanında farkındalık programları düzenlenmelidir. Bu programlar, sektörde çalışan profesyonellerin bilgi ve becerilerini güncellemelerine yardımcı olacak ve değişen siber güvenlik ekosistemine uyum sağlamalarını kolaylaştıracaktır.

- Bilgi Teknolojileri ve İletişim Kurumu (BTK) gibi devlet kurumları, üniversitelerle iş birliği yaparak öğrencilere staj veya laboratuvar imkanı sunmalıdır. Bu programlar, öğrencilerin devletin siber güvenlik politikalarını ve uygulamalarını yerinde görmelerini ve katkıda bulunmalarını sağlamalıdır. Ayrıca, bu programlar süresince öğrencilerden belirli yapay zeka projeleri üzerinde çalışmaları ve bu projelerin sonuçlarını ve çıktılarını raporlamaları istenmelidir. BTK Akademi de yapay zeka ve makine öğrenmesi ile ilgili daha kapsamlı eğitimler eklenebilir ve bunun için üniversitelerle iş birliği yapılabilir.

Teknoloji ve Altyapı

Dünya örneklerinde incelenen makine öğrenmesi projelerinin ülkemizde de hayata geçirilebilmesi için gerekli çalışmalar ve teşvikler gerçekleştirilmektedir. Ancak, anlaşılacağı üzere makine öğrenmesi modellerinin etkinliği, kullanılan veri setlerinin kalitesine ve miktarına bağlıdır. Bu nedenle, kaliteli ve etiketlenmiş veri setlerine erişim sağlanması verimli ve yüksek kaliteli modeller eğitilmesini sağlayacaktır.

- BTK öncülüğünde, kamu kurumları ve özel sektör arasında veri paylaşımını kolaylaştıracak güvenli bir platform oluşturulmalıdır. Bu platform, siber güvenlik verilerinin anonimleştirilerek milli güvenliği arttırmak için gerçekleştirilecek projelerde kullanılmalıdır. Platformun kullanıcıları için veri güvenliği ve gizliliği protokolleri oluşturulmalı ve bu protokoller sıkı bir şekilde takip edilip uygulanmalıdır.
- Kamu kurumları ve özel sektör, veri setlerinin etiketlenmesi için ortak projeler başlatmalıdır. Bu projelerde, yapay zeka ve makine öğrenmesi uzmanları, veri bilimcileri, analistleri, madencileri ve siber güvenlik uzmanları birlikte çalışarak veri setlerini etiketlemelidir.

Makine öğrenmesi algoritmaları, siber güvenlik alanında spesifik ihtiyaçlara göre optimize edilmelidir. Özellikle anomali tespiti, zararlı yazılım analizi gibi alanlarda özel algoritmalar geliştirilmelidir. Ülkemizde ağ izleme ve oltalama tespiti gibi alanlarda çalışmalar bulunmaktadır. Özel sektörün de makine öğrenmesini kullanan siber güvenlik projeleri üstüne çalıştığı anlaşılmaktadır fakat yeteri sayıda son kullanıcı ürünü bulunmamaktadır. Özellikle; anomali tespiti ve zararlı yazılım analizi çalışılması gereken konular olarak karşımıza çıkmaktadır. Bu çözümler geliştirildikten sonra bulut tabanlı hizmetlerle desteklenmelidir. İncelemelerde ülkelerin hizmetlerini bulut bilişim çözümleri ile entegre ettikleri anlaşılmaktadır. Bulut bilişim, büyük veri setlerinin işlenmesi ve depolanması için altyapı sunmaktadır. Ayrıca, bulut tabanlı çözümler, siber güvenlik operasyonlarının ölçeklenebilirliğini artırmakta ve kurumlar için maliyetleri düşürmektedir. Ülkemizdeki siber güvenlik firmaları ve kamu kurumları, bulut tabanlı çözümleri de göz önünde bulundurmalıdır. Bulut bilişim

çözümünde de yerli ve milli ürünler tercih edilmeli ve verilerin yurtdışına çıkmaması için gerekli önlemler alınmalıdır.

Ülkemizde yeterince faydalanılmayan başka bir makine öğrenmesi uygulaması ise makine öğrenmesi tabanlı tehdit istihbaratı sistemleridir. Bu sistemlerin temelinde makine öğrenmesinin yanı sıra doğal dil işleme yöntemlerinden faydalanıldığı tespit edilmiştir. Ülke incelemelerinden anlaşılacağı üzere, bu sistemler, gerçek zamanlı olarak tehditleri tespit etmek ve bazı durumlarda da otomatik olarak karşılık vermek üzere kullanılmaktadır. Özellikle, siber tehditlerin proaktif bir şekilde tespit edilmesi ve önlenmesi için makine öğrenmesi modelleri kullanılmalıdır. Bu modeller, sürekli olarak güncellenerek yeni ortaya çıkan tehditlere karşı etkin bir savunma sağlayabilmektedir. Makine öğrenmesi kullanarak uyarıların önceliklendirilmesi ve yönetilmesi de sağlanmaktadır. Böylece, siber güvenlik uzmanlarının daha kritik olaylara odaklanabilmesi ve yanlış pozitif uyarıların sayısının azaltılması sağlanmaktadır. Bu sistemler, siber güvenlik operasyon merkezleri tarafından kullanılarak insan kaynağından daha verimli bir şekilde yararlanılması sağlanabilir.

- Türkiye'deki üniversiteler ve araştırma merkezleri, Türkçe dil işleme projelerine odaklanmalıdır. Bu projeler, metin analizi, spam tespiti, duygu analizi gibi alanlarda kullanılabilecek Türkçe NLP modelleri geliştirmelidir. TÜBİTAK ve Sanayi ve Teknoloji Bakanlığı, bu projelere fon veya teşvik sağlayarak yerli dil işleme teknolojilerinin geliştirilmesini desteklemelidir.

Araştırma ve Geliştirme

Ülkemizin ihtiyaçlarına uygun yerli ve milli makine öğrenmesi çözümleri geliştirilmelidir. Bu çözümler, Türkçe dil işleme ve yerel tehditlere özel algoritmalar içerecek şekilde eğitilebilir. Yerli ve milli çözümler, ülkemizin siber güvenlik alanındaki tam bağımsızlığını sağlamasından rol oynayacak ve dışa bağımlılığımızı azaltacaktır. Ayrıca, yerli ve milli yazılım ve donanım geliştirme faaliyetleri uzun vadede ülkemizin ekonomik gücüne katkıda bulunacaktır.

Uluslararası araştırma ve geliştirme projelerine katılım sağlanmalı ve bu alanda bilgi alışverişi de yapılmalıdır. Bu iş birlikleri ile daha gelişmiş ve yenilikçi çözümlerin ülkemize kazandırılması sağlanacaktır. Özellikle, Amerika Birleşik Devletleri diğer uluslararası ya da ulusal kuruluşlarla ortak projeler yürütmektedir. Ülkemizin de benzeri girişimlere dahil olması ülkemize hızlı bir şekilde bilgi aktarılmasını sağlayabilir. Bu projeler, ülkemizin siber güvenlik kapasitesini artıracak ve küresel düzeyde rekabet gücümüzü yükseltecektir. Üniversiteler ve sanayi kuruluşları arasında iş birliği teşvik edilmelidir. Bu iş birliği sayesinde, akademik araştırmaların pratik uygulamalara dönüştürülmesi sağlanacaktır. Ayrıca, üniversitelerde yapılan araştırmalar, sanayi kuruluşlarının ihtiyaçlarına yönelik çözümler sunabilmektedir. Bu iş birliği, ülkemizdeki siber güvenlik ekosisteminin gelişmesine katkı sağlayarak gençlerimizin bilgi birikimlerini sektöre daha verimli bir biçimde entegre etmesine olanak sağlayacaktır.

Makine öğrenmesi, siber güvenlik alanında etkili bir araçtır ve ülkemiz bu teknolojiyi etkin bir şekilde kullanabilmek için gerekli kapasiteye sahiptir. Devletimizin MGK açıklaması gibi yaptığı açıklamalar ve sektörde önerilen çalışmalar yapay zekaya ne kadar önem verildiğini de göstermektedir. Eğitimden teknolojiye, politikalardan uygulamaya kadar geniş bir yelpazede yapılacak hamleler ile ülkemizin makine öğrenmesi kullanarak siber güvenlik alanında daha güçlü ve dirençli bir yapıya kavuşması sağlanabilecektir.

Yasal ve Mevzuat Düzenlemeleri

BTK tarafından haberleşme sektöründe yapay zekanın daha etkili ve güvenli bir şekilde kullanılmasını sağlamak ve ilgili denetim mekanizmalarını oluşturmak için yürütmeyi temin edecek bir düzenleme oluşturulmalı ve BTK'nın yapay zekanın yasal düzenlemeleri konusunda kendi sektörünü düzenleyen ilk kurum olarak ülkemizde öncülük etmesi sağlanmalıdır. Bu düzenleme ile yapay zekanın haberleşme sektöründe kullanılması durumunda hangi alanlarda kullanıldığına bağlı olarak nasıl yaptırımlara tabi olacağı ve yapay zeka eğitilmesi için gerekli olan verilerin temini ve güvenliği gibi hususlar sektörün ihtiyaçlarına uygun bir şekilde düzenlenmelidir. Bu düzenleme

özellikle makine öğrenmesi modellerinin geliştirilmesi için kullanılan veri setlerinin güvenliği ve doğru amaca uygun kullanımı ile ilgili hükümler içermelidir. Bunun yanı sıra, yapay zekanın karar verme süreçlerinde şeffaflık, hesap verebilirlik ve etik ilkeler gibi temel prensiplerin gözetilmesi gerektiği vurgulanmalı; bu doğrultuda yapay zeka sistemlerinin denetlenebilirliği ve izlenebilirliğini sağlamak için teknik standartların oluşturulması önerilmelidir.

Bu düzenleme 5809 Sayılı Elektronik Haberleşme Kanunu'nun Madde 12, 2. Fıkrası;

> *"Kurum, işletmecilere sektörün ihtiyaçları, uluslararası düzenlemeler, teknolojide meydana gelen gelişmeler gibi hususları gözeterek aşağıdaki hususlar başta olmak üzere, mevzuat doğrultusunda yükümlülükler getirebilir:*
> *d) Kişisel veri ve gizliliğin korunması.*
> *e) Tüketicinin korunması.*
> *f) Kuruma bilgi ve belge verilmesi."*

tarafından BTK'ya verilen yetkiler doğrultusunda hazırlanabilir Ayrıca, bu düzenleme hazırlanırken ulusal güvenlik boyutları da dikkate alınmalı ve yapay zekanın potansiyel kötüye kullanım senaryolarına karşı tedbirlerin tanımlanması hedeflenmelidir. Böylelikle, haberleşme altyapısının güvenliği ve sürekliliği konusunda bir standart oluşturulabilecektir. Bilgi Teknolojileri ve iletişim Kurumu Teşkilat Yönetmeliğinin 27. Maddesi'nin 1. Fıkrası;

> *"e) Bilgi teknolojilerinin ve elektronik uygulamaların gelişimini izlemek, gelişmeler doğrultusunda çalışmalar yapmak ve gerekmesi halinde ülkemizde yapılması gereken iş ve işlemlere ilişkin hukuki, teknik ve idarî konularda öneriler geliştirmek.*
> *f) İlgili mevzuatı çerçevesinde bilgi toplumuna dönüşüm amacıyla gerçekleştirilen çalışmaları izlemek, bu çalışmalara katkıda bulunmak, bu çalışmalar doğrultusunda Kurum bünyesinde ve sektör nezdinde yapılması gereken çalışmaları koordine etmek."*

ifadeleri doğrultusunda Bilgi Teknolojileri Dairesi Başkanlığı tarafından uluslararası örnekler incelenerek söz konusu yönetmeliğin hazırlanmasına katkı sağlanabilir. Bu kapsamda, uluslararası kuruluşlar ve standart belirleyici organizasyonlarla iş birliği yapılması ve bu düzenlemenin Türkiye'nin uluslararası alandaki rekabet gücünü artıracak şekilde yapılandırılması hedeflenmelidir.

Makine Öğrenmesi Tabanlı Anomali Tespit ve Müdahale Sistemi

USOM tarafından uygun görülen kritik altyapılarda ve kamu kurumlarında siber güvenlik tehditlerini erken tespit etmek ve etkili bir şekilde müdahale etmek amacıyla "Makine Öğrenmesi Tabanlı Anomali Tespit ve Müdahale Sistemi" projesi geliştirilebilir. Böylece önem verilen sistemlerde yaşanan olaylar anlık olarak tespit edilip olay müdahale süreçlerinde çok büyük bir zaman kazanılabilir.

Makine öğrenmesi ve yapay zeka tekniklerini kullanarak ağ trafiğindeki anormallikleri ve potansiyel siber saldırıları gerçek zamanlı olarak tespit etmesi üzere eğitilecek bu sistem, özellikle kamu kurumları ve kritik altyapılarda kullanılmak üzere teknik altyapısı USOM tarafından sağlanarak bir veri merkezi üzerinden bulut bilişim modeli kullanılarak hizmet verebilir.

Sistem, ağ trafiği, sistem günlükleri ve kullanıcı davranışları gibi çeşitli veri kaynaklarından sürekli olarak veri toplayarak ve bu verileri ön işleme sürecinden geçirerek analiz için uygun hale getirecektir.

Denetimli ve denetimsiz öğrenme algoritmaları ile eğitilen makine öğrenmesi modelleri, geçmişteki siber saldırı verilerini yeni tehditleri tespit etmekte kullanacaktır. Derin öğrenme modelleri ise daha karmaşık saldırı senaryolarını tespit edebilmek için kullanılabilir.

Sistem, tespit edilen anomali ve tehditlere karşı otomatik olarak tepki veren bir müdahale mekanizması geliştirerek tehdit düzeyine göre farklı müdahale stratejileri uygulanmasına olanak sağlayacaktır. Ayrıca, anomali ve tehdit tespitlerinin görsel olarak izlenebilmesi için kullanıcı dostu bir arayüz geliştirilerek ve sistem yöneticilerine düzenli raporlar ve anlık uyarılar sunarak karar alma süreçlerinin desteklenmesi sağlanabilir.

Toplanan verilerin güvenli bir şekilde saklanması ve işlenmesi ile gizlilik ilkelerine uygun olarak kişisel verilerin korunması da gerekmektedir. Bu proje, siber saldırıları

erken aşamada tespit ederek zarar vermeden müdahale edilmesinin sağlanması, otomatik müdahale mekanizması ile insan müdahalesini en aza indirerek tehditlere karşı önlem alınabilmesi ve güvenlik ekiplerinin yükünü azaltarak daha verimli çalışmalarını sağlayacaktır.

Sürekli öğrenen ve güncellenen sistem sayesinde yeni ve gelişen tehditlere karşı proaktif bir savunma oluşturulup, USOM'un siber güvenlik kapasitesini artırarak kamu kurumları ve kritik altyapılarda güvenliği sağlanmasına katkı sağlayacaktır.

KAYNAKLAR

Airbus. (2024, June 1). *Our History*. Https://Www.Airbus.Com/En/Our-History.

Airbus CyberSecurity. (2024a, June 2). *About Us*. Https://Www.Cyber.Airbus.Com/Information-Centre/about-Us/.

Airbus CyberSecurity. (2024b, June 2). *Innovation-Cognification*. Https://Www.Cyber.Airbus.Com/Innovation/#cognification.

Airbus CyberSecurity. (2024c, June 23). *Automating the SOC – Towards AI-Based Incident Response in the Factory of the Future*. Https://Www.Cyber.Airbus.Com/Automating-the-Soc-towards-Ai-Based-Incident-Response-in-the-Factory-of-the-Future-2/.

Aiyer, B., Caso, J., Russell, P., & Sorel, M. (2022). *October 2022 Cybersecurity Practice*.

Alhogail, A., & Alsabih, A. (2021). Applying machine learning and natural language processing to detect phishing email. *Computers and Security*, *110*. https://doi.org/10.1016/j.cose.2021.102414

Almukaynizi, M., Nunes, E., Dharaiya, K., Senguttuvan, M., Shakarian, J., & Shakarian, P. (2017). Proactive identification of exploits in the wild through vulnerability mentions online. *2017 International Conference on Cyber Conflict (CyCon U.S.)*, 82–88. https://doi.org/10.1109/CYCONUS.2017.8167501

Alrawili, R., Oliva, M., Honnef, A., Sawall, E., & Alqahtani, A. A. S. (2022). Malware and Average Individual. *APWiMob 2022 - Proceedings: 2022 IEEE Asia Pacific Conference on Wireless and Mobile*. https://doi.org/10.1109/APWiMob56856.2022.10014080

Amazon. (2024a, May 1). *Amazon Macie*. Https://Aws.Amazon.Com/Tr/Macie/.

Amazon. (2024b, May 15). *About Guavus*. Https://Docs.Aws.Amazon.Com/Whitepapers/Latest/Guavus-5G-Iq-Nwdaf-on-Aws/about-Guavus.Html.

Amer, E., & Zelinka, I. (2020). A dynamic Windows malware detection and prediction method based on contextual understanding of API call sequence. *Computers and Security*, *92*. https://doi.org/10.1016/j.cose.2020.101760

Andresini, G., Pendlebury, F., Pierazzi, F., Loglisci, C., Appice, A., & Cavallaro, L. (2021). INSOMNIA. *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security*, 111–122. https://doi.org/10.1145/3474369.3486864

Andress, J., & Winterfeld, S. (2014). Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners. In *Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners*. https://doi.org/10.1016/B978-0-12-375067-9.09978-2

ANSSI. (2024a, May 16). *About French Cybersecurity Agency (ANSSI)*. Https://Cyber.Gouv.Fr/En/about-French-Cybersecurity-Agency-Anssi.

ANSSI. (2024b, May 17). *ANSSI-FR / SecuML*. Https://Github.Com/ANSSI-FR/SecuML?Tab=readme-Ov-File.

Apruzzese, G., Andreolini, M., Marchetti, M., Venturi, A., & Colajanni, M. (2020). Deep Reinforcement Adversarial Learning Against Botnet Evasion Attacks. *IEEE Transactions on Network and Service Management*, *17*(4), 1975–1987. https://doi.org/10.1109/TNSM.2020.3031843

Apruzzese, G., Colajanni, M., & Marchetti, M. (2019). Evaluating the effectiveness of Adversarial Attacks against Botnet Detectors. *2019 IEEE 18th*

International Symposium on Network Computing and Applications (NCA), 1–8. https://doi.org/10.1109/NCA.2019.8935039

Apruzzese, G., Laskov, P., & Tastemirova, A. (2022). SoK: The Impact of Unlabelled Data in Cyberthreat Detection. *2022 IEEE 7th European Symposium on Security and Privacy (EuroS&P)*, 20–42. https://doi.org/10.1109/EuroSP53844.2022.00010

Apruzzese, G., Marchetti, M., Colajanni, M., Zoccoli, G. G., & Guido, A. (2017). Identifying malicious hosts involved in periodic communications. *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)*, 1–8. https://doi.org/10.1109/NCA.2017.8171326

Arf, C. (1959). Makineler Düşünebilir mi ve Nasıl Düşünebilir? *Atatürk Üniversitesi 1958- 1959 Öğretim Yılı Halk Konferansları*.

Balas, V. E., Kumar, R., & Srivastava, R. (2020). *Recent Trends and Advances in Artificial Intelligence and Internet of Things* (V. E. Balas, R. Kumar, & R. Srivastava, Eds.; Vol. 172). Springer International Publishing. https://doi.org/10.1007/978-3-030-32644-9

BEREC. (2024). *BEREC high-level position on artificial intelligence and virtual worlds*.

Biggio, B., & Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. *Pattern Recognition*, *84*, 317–331. https://doi.org/10.1016/j.patcog.2018.07.023

Bilge, L., Sen, S., Balzarotti, D., Kirda, E., & Kruegel, C. (2014). Exposure: A Passive DNS Analysis Service to Detect and Report Malicious Domains. *ACM Transactions on Information and System Security*, *16*(4), 1–28. https://doi.org/10.1145/2584679

Blackberry. (2024, May 6). *Bad Rabbit Ransomware*. Https://Www.Blackberry.Com/Us/En/Solutions/Endpoint-Security/Ransomware-Protection/Bad-Rabbit.

Bougourzi, H., & Bougourzi, A. H. (2020). *Understanding the mathematics of artificial neural networks*. https://www.researchgate.net/publication/350811538

Buczak, A. L., & Guven, E. (2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys and Tutorials*, *18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

Chakraborty, T., Pierazzi, F., & Subrahmanian, V. S. (2020). EC2: Ensemble Clustering and Classification for Predicting Android Malware Families. *IEEE Transactions on Dependable and Secure Computing*, *17*(2), 262–277. https://doi.org/10.1109/TDSC.2017.2739145

Chaudhary, S., O'Brien, A., & Xu, S. (2020). Automated Post-Breach Penetration Testing through Reinforcement Learning. *2020 IEEE Conference on Communications and Network Security (CNS)*, 1–2. https://doi.org/10.1109/CNS48642.2020.9162301

Chen, H., Liu, J., Liu, R., Park, N., & Subrahmanian, V. S. (2019). VASE: A Twitter-Based Vulnerability Analysis and Score Engine. *2019 IEEE International Conference on Data Mining (ICDM)*, *2019-November*, 976–981. https://doi.org/10.1109/ICDM.2019.00110

Chua, Z. L., Shen, S., Saxena, P., & Liang, Z. (2017). Neural Nets Can Learn Function Type Signatures From Binaries. *26th USENIX Security Symposium (USENIX Security 17)*, 99–116. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/chua

CISA. (2024a, May 1). *About CISA*. Https://Www.Cisa.Gov/About.

CISA. (2024b, May 1). *CISA Artificial Intelligence Use Cases*. Https://Www.Cisa.Gov/Ai/Cisa-Use-Cases.

Cloudflare. (2024a, May 3). *What was the WannaCry ransomware attack?* Https://Www.Cloudflare.Com/Learning/Security/Ransomware/Wannacry-Ransomware/.

Cloudflare. (2024b, May 5). *What are Petya and NotPetya?* Https://Www.Cloudflare.Com/Learning/Security/Ransomware/Petya-Notpetya-Ransomware/.

Communications Security Establishment (Canada). (2022). *An introduction to the cyber threat environment.*

Computer Security Division. (2006). *FIPS PUB 200 Minimum Security Requirements for Federal Information and Information Systems*.

Computer Security Division. (2011). *Managing information security risk*. https://doi.org/10.6028/NIST.SP.800-39

Darktrace. (2024). *DARKTRACE ACTIVEAI SECURITY PLATFORM*.

DDO. (2021). *Ulusal Yapay Zekâ Stratejisi (UYZS) 2021-2025*.

Deep Instict. (2024). *DIANNA-Deep Instinct's Artificial Neural Network Assistant Analyzes Threat Delivery Files*. www.deepinstinct.com

Deep Instinct. (2024a, May 1). *Deep Instinct About Us*. Https://Www.Deepinstinct.Com/about-Us.

Deep Instinct. (2024b, May 2). *Deep Instinct Prevention Platform.* Https://Www.Deepinstinct.Com/Platform.

Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). *Mathematics for Machine Learning*. Cambridge University Press. https://doi.org/10.1017/9781108679930

Diale, M., Celik, T., & Van Der Walt, C. (2019). Unsupervised feature learning for spam email filtering. *Computers and Electrical Engineering*, *74*, 89–104. https://doi.org/10.1016/j.compeleceng.2019.01.004

Du, M., Li, F., Zheng, G., & Srikumar, V. (2017). DeepLog: Anomaly detection and diagnosis from system logs through deep learning. *Proceedings of the ACM Conference on Computer and Communications Security*, 1285–1298. https://doi.org/10.1145/3133956.3134015

Dundar, M., Krishnapuram, B., Bi, J., & Rao, R. (2007). Learning Classifiers When the Training Data Is Not IID. *IJCAI International Joint Conference on Artificial Intelligence*, 756–761.

ENISA. (2021). *ENISA THREAT LANDSCAPE FOR SUPPLY CHAIN ATTACKS JULY 2021 ENISA THREAT LANDSCAPE FOR SUPPLY CHAIN ATTACKS ABOUT ENISA*. https://doi.org/10.2824/168593

Eprs, & Rapporteurs. (2024). *BRIEFING EU Legislation in Progress Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain Union legislative acts Committees responsible*.

European Commission. (2020). *On Artificial Intelligence-A European approach to excellence and trust White Paper on Artificial Intelligence A European approach to excellence and trust*. https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf.

European Parliament. (2024). *OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of … laying down harmonised rules on artificial intelligence and amending Regulations (EC) No (Artificial Intelligence Act) (Text with EEA relevance)*.

European Union. (2024, May 23). *About AI4EU*. Https://Www.Ai4europe.Eu/about-Ai4eu.

Fischer, E. A. (2016). *Cybersecurity Issues and Challenges: In Brief*. www.crs.gov

GCHQ. (2024, May 14). *Mission Overview*. Https://Www.Gchq.Gov.Uk/Section/Mission/Overview.

Ghanem, M. C., & Chen, T. M. (2018). Reinforcement Learning for Intelligent Penetration Testing. *2018 Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, 185–192. https://doi.org/10.1109/WorldS4.2018.8611595

Hatfield, J. M. (2018). Social engineering in cybersecurity: The evolution of a concept. *Computers and Security*, *73*, 102–113. https://doi.org/10.1016/j.cose.2017.10.008

Haykin, S. S. (2009). *Neural networks and learning machines*. Prentice Hall/Pearson.

Hochreiter, S., & Schmidhuber, J. (1997). Long-Short Term Mermory. *Neural Computation 9(8):1735-1780.*

Hu, X., Bhatkar, S., Griffin, K., & Shin, K. G. (2013). MutantX-S: Scalable Malware Clustering Based on Static Features. *USENIX ATC'13: Proceedings of the 2013 USENIX Conference on Annual Technical Conference*, 187–198.

Inria. (2024a, May 18). *Machine Learning Network System Security*. Https://Team.Inria.Fr/Mlns2/.

Inria. (2024b, May 21). *Digital Security*. Https://Www.Inria.Fr/En/Digital-Security.

ITU. (2024a, June 8). *ITU Home Page*. Https://Www.Itu.Int/En/Pages/Default.Aspx.

ITU. (2024b, June 12). *AI for Good*. Https://Aiforgood.Itu.Int/.

Joe Turner, A., Editorial Board, U., Meyer, B., Zurich, E., Rannenberg, K., & Bramer, M. A. (2010). *IFIP Advances in Information and Communication Technology 330 Editor-in-Chief Security and Privacy Protection in Information Processing Systems*.

Jordaney, R., Sharad, K., Dash, S. K., Wang, Z., Papini, D., Nouretdinov, I., & Cavallaro, L. (2017). Transcend: Detecting Concept Drift in Malware Classification Models. *26th USENIX Security Symposium (USENIX Security 17)*, 625–642. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/jordaney

Kang, C., Park, N., Prakash, B. A., Serra, E., & Subrahmanian, V. S. (2016). Ensemble Models for Data-driven Prediction of Malware Infections. *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, 583–592. https://doi.org/10.1145/2835776.2835834

Karim, A., Azam, S., Shanmugam, B., Kannoorpatti, K., & Alazab, M. (2019). A Comprehensive Survey for Intelligent Spam Email Detection. *IEEE Access*, *7*, 168261–168295. https://doi.org/10.1109/ACCESS.2019.2954791

Kettani, H., & Wainwright, P. (2019). On the Top Threats to Cyber Systems. *2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT)*, 175–179. https://doi.org/10.1109/INFOCT.2019.8711324

Kharraz Amin, Robertson William, Balzarotti Davide, Bilge Leyla, & Kirda Engin. (2015). *Detection of Intrusions and Malware, and Vulnerability Assessment* (M. Almgren, V. Gulisano, & F. Maggi, Eds.; Vol. 9148). Springer International Publishing. https://doi.org/10.1007/978-3-319-20550-2

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, *37*(2), 233–243. https://doi.org/10.1002/aic.690370209

Kshetri, N. (2021). Economics of Artificial Intelligence in Cybersecurity. *IT Professional*, *23*(5), 73–77. https://doi.org/10.1109/MITP.2021.3100177

Lancaster, E., Chakraborty, T., & Subrahmanian, V. S. (2018). MALTP: Parallel Prediction of Malicious Tweets. *IEEE Transactions on Computational Social Systems*, *5*(4), 1096–1108. https://doi.org/10.1109/TCSS.2018.2869171

Lau, F., Rubin, S. H., Smith, M. H., & Trajkovic, L. (2000). *Distributed Denial of Service Attacks*.

Lecun, Y., Haffner, P., Bottou, L., & Bengio, Y. (1999). *Object Recognition with Gradient-Based Learning*. http://www.research.att.com/~yann

Lella, I., Theocharidou, M., Tsekmezoglou, E., Svetozarov Naydenov, R., Ciobanu, C., Malatras, A., & European Union Agency for Cybersecurity. (2022). *ENISA threat landscape 2022 : July 2021 to July 2022.*

Levy, I. (2019). *Active Cyber Defence - The Second Year*.

Liang, B., Su, M., You, W., Shi, W., & Yang, G. (2016a). Cracking Classifiers for Evasion. *Proceedings of the 25th International Conference on World Wide Web*, 345–356. https://doi.org/10.1145/2872427.2883060

Liang, B., Su, M., You, W., Shi, W., & Yang, G. (2016b). Cracking Classifiers for Evasion. *Proceedings of the 25th International Conference on World Wide Web*, 345–356. https://doi.org/10.1145/2872427.2883060

Lopes, N., & Ribeiro, B. (2015). Deep Belief Networks (DBNs). In *Studies in Big Data* (Vol. 7, pp. 155–186). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-319-06938-8_8

Lucci, S., Musa, S. M., & Kopec, D. (2022). *Artificial Intelligence in the 21 st Century Third Edition*.

Martínez Garre, J. T., Gil Pérez, M., & Ruiz-Martínez, A. (2021). A novel Machine Learning-based approach for the detection of SSH botnet infection. *Future Generation Computer Systems*, *115*, 387–396. https://doi.org/10.1016/j.future.2020.09.004

McElwee, S., Heaton, J., Fraley, J., & Cannady, J. (2017). Deep learning for prioritizing and responding to intrusion detection alerts. *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, 1–5. https://doi.org/10.1109/MILCOM.2017.8170757

McKinnel, D. R., Dargahi, T., Dehghantanha, A., & Choo, K.-K. R. (2019). A systematic literature review and meta-analysis on artificial intelligence in penetration testing and vulnerability assessment. *Computers & Electrical Engineering*, *75*, 175–188. https://doi.org/10.1016/j.compeleceng.2019.02.022

Meskauskas, T. (2024, May 2). *Spora ransomware virus – removal and decryption options*. Https://Www.Pcrisk.Com/Removal-Guides/10824-Spora-Ransomware.

MGK. (2024, May 28). *28 Mayıs 2024 Tarihli Toplantı*. Https://Www.Mgk.Gov.Tr/Index.Php/28-Mayis-2024-Tarihli-Toplanti.

Nadeem, A., Verwer, S., Moskal, S., & Yang, S. J. (2021). Alert-driven Attack Graph Generation using S-PDFA. *IEEE Transactions on Dependable and Secure Computing*, *19*(2), 1–1. https://doi.org/10.1109/TDSC.2021.3117348

NATO. (2024, June 13). *Summary of the NATO Artificial Intelligence Strategy*. Https://Www.Nato.Int/Cps/En/Natohq/Official_texts_187617.Htm.

NCSC. (2020). *Active Cyber Defence - The Third Year*. www.ncsc.gov.uk/section/products-services/active-cyber-defence

NCSC. (2022). *Active Cyber Defence - The Fifth Year*.

NCSC. (2023). *Active Cyber Defence-The Sixth Year*.

NCSC. (2024, May 13). *About the NCSC*. Https://Www.Ncsc.Gov.Uk/Section/about-Ncsc/What-We-Do.

Needham, R. M. (1993). *DENIAL OF SERVICE*.

OECD. (2024, June 6). *OECD AI Principles overview*. Https://Oecd.Ai/En/Ai-Principles.

Okutan, A., & Yang, S. J. (2019). ASSERT: attack synthesis and separation with entropy redistribution towards predictive cyber defense. *Cybersecurity*, *2*(1). https://doi.org/10.1186/s42400-019-0032-0

Orange. (2024, May 30). *Security Event Intelligence*. Https://Www.Orange-Business.Com/En/Solutions/Security/Security-Event-Intelligence.

Pacepa, I. M., & Rychlak, R. J. (2003). *DISINFORMATION FORMER SPY CHIEF REVEALS SECRET STRATEGIES FOR UNDERMINING FREEDOM, ATTACKING RELIGION, AND PROMOTING TERRORISM ION MIHAI PACEPA AND Prof. RONALD J. RYCHLAK*.

PaloAltoNetworks. (2023). *Cortex XDR*.

PaloAltoNetworks. (2024, May 23). *What Is DNS Tunneling?* Https://Www.Paloaltonetworks.Com/Cyberpedia/What-Is-Dns-Tunneling.

Paquet-Clouston, M., Haslhofer, B., & Dupont, B. (2019). Ransomware payments in the Bitcoin ecosystem. *Journal of Cybersecurity*, *5*(1), 1–11. https://doi.org/10.1093/cybsec/tyz003

Parmar, R. (2024, March 30). *Training Deep Neural Networks*. Towards Data Science.

Pendlebury, F., Pierazzi, F., Jordaney, R., Kinder, J., & Cavallaro, L. (2018). *TESSERACT: Eliminating Experimental Bias in Malware Classification across Space and Time*. https://doi.org/10.48550/arXiv.1807.07838

Pierazzi, F., Pendlebury, F., Cortellazzi, J., & Cavallaro, L. (2019). *Intriguing Properties of Adversarial ML Attacks in the Problem Space*. http://arxiv.org/abs/1911.02142

Portnoff, R. S., Afroz, S., Durrett, G., Kummerfeld, J. K., Berg-Kirkpatrick, T., McCoy, D., Levchenko, K., & Paxson, V. (2017). Tools for Automated Analysis of Cybercriminal Markets. *Proceedings of the 26th International Conference on World Wide Web*, 657–666. https://doi.org/10.1145/3038912.3052600

Prasad, R., & Rohokale, V. (2015). *Cyber Security: Analytics, Technology and Automation* (M. Lehto & P. Neittaanmäki, Eds.; Vol. 78). Springer International Publishing. https://doi.org/10.1007/978-3-319-18302-2

R., V., Alazab, M., Jolfaei, A., K.P., S., & Poornachandran, P. (2019). Ransomware Triage Using Deep Learning: Twitter as a Case Study. *2019 Cybersecurity and Cyberforensics Conference (CCC)*, 67–73. https://doi.org/10.1109/CCC.2019.000-7

Ramanathan, V., Wang, R., & Mahajan, D. (2021). PreDet: Large-scale weakly supervised pre-training for detection. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2845–2855. https://doi.org/10.1109/ICCV48922.2021.00286

Rannenberg, K., Varadharajan, V., & Weber, C. (2010). The 5 Waves of Information Security-From Kristian Beckman to the Present. In *IFIP AICT* (Vol. 330). www.ifip.org

Rawindran, N., Jayal, A., & Prakash, E. (2021). *Artificial Intelligence and Machine Learning within the context of Cyber Security used in the UK SME sector*.

Razaulla, S., Fachkha, C., Markarian, C., Gawanmeh, A., Mansoor, W., Fung, B. C. M., & Assi, C. (2023). The Age of Ransomware: A Survey on the Evolution, Taxonomy, and Research Directions. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3268535

Rieck, K., Holz, T., Willems, C., Düssel, P., & Laskov, P. (2008). Learning and Classification of Malware Behavior. In *Detection of Intrusions and Malware, and Vulnerability Assessment* (pp. 108–125). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70542-0_6

Robert Keim. (2020, February 5). *Incorporating Bias Nodes Into Your Neural Network*. Https://Www.Allaboutcircuits.Com/Technical-Articles/Incorporating-Bias-Nodes-into-Your-Neural-Network/#:~:Text=Bias%20nodes%2C%20which%20can%20be,Is%20chosen%20by%20the%20designer.

Rokach, L., Maimon, O., & Shmueli, E. (2023). Machine Learning for Data Science Handbook. In L. Rokach, O. Maimon, & E. Shmueli (Eds.), *Machine Learning for Data Science Handbook*. Springer International Publishing. https://doi.org/10.1007/978-3-031-24628-9

Rostamizadeh, A., & Talwalkar, A. (2012). *Foundations of Machine Learning second edition*.

Rusk, N. (2015). *Deep learning*. https://doi.org/10.1101/028399

Sak, H., Senior, A., & Beaufays, F. (2014). *Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition*. http://arxiv.org/abs/1402.1128

Sapienza, A., Bessi, A., Damodaran, S., Shakarian, P., Lerman, K., & Ferrara, E. (2017). Early Warnings of Cyber Threats in Online Discussions. *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, 667–674. https://doi.org/10.1109/ICDMW.2017.94

Sector, D. (2022). *Measuring digital development Facts and Figures 2022*.

Sharafaldin, I., Habibi Lashkari, A., & Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, *2018-January*, 108–116. https://doi.org/10.5220/0006639801080116

Sipola, T., Kokkonen, T., & Karjalainenn, M. (2023). *Artificial Intelligence and Cybersecurity* (T. Sipola, T. Kokkonen, & M. Karjalainen, Eds.). Springer International Publishing. https://doi.org/10.1007/978-3-031-15030-2

SparkCognition. (2018). *DeepArmor Platform Architecture*.

Srndic, N., & Laskov, P. (2014a). Practical Evasion of a Learning-Based Classifier: A Case Study. *2014 IEEE Symposium on Security and Privacy*, 197–211. https://doi.org/10.1109/SP.2014.20

Srndic, N., & Laskov, P. (2014b). Practical Evasion of a Learning-Based Classifier: A Case Study. *2014 IEEE Symposium on Security and Privacy*, 197–211. https://doi.org/10.1109/SP.2014.20

SSYZ. (2024, May 23). *SSYZ Hakkında*. Https://Ssyz.Org.Tr/About.

STM. (2024, May 27). *Siber Güvenlikte Yapay Zekâ*. Https://Www.Stm.Com.Tr/Tr/Inovasyon/Siber-Guvenlikte-Yapay-Zeka.

Study Group, I. (2008). *ITU-T Rec. X.1205 (04/2008) Overview of cybersecurity*.

Su, Y. H., Cho, M. C. Y., & Huang, H. C. (2019). False alert buster: An adaptive approach for NIDS false alert filtering. *ACM International Conference Proceeding Series*, 58–62. https://doi.org/10.1145/3366650.3366657

Sweet, C., Moskal, S., & Yang, S. J. (2020). On the Variety and Veracity of Cyber Intrusion Alerts Synthesized by Generative Adversarial Networks. *ACM Transactions on Management Information Systems*, *11*(4), 1–21. https://doi.org/10.1145/3394503

Taylor, J. G. (John G. (1993). *The promise of neural networks*. Springer-Verlag.

TDK. (2023, August 25). *Türk Dil Kurumu Sözlükleri*. Türk Dil Kurumu.

Thales. (2024, May 29). *Thales Group History*. Https://Www.Thalesgroup.Com/En/Global/Group/History.

Tian, K., Jan, S. T. K., Hu, H., Yao, D., & Wang, G. (2018). Needle in a haystack: Tracking down elite phishing domains in the wild. *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, 429–442. https://doi.org/10.1145/3278532.3278569

UAB. (2020). *ULUSAL SİBER GÜVENLİK STRATEJİSİ*. https://hgm.uab.gov.tr/uploads/pages/siber-guvenlik/ulusal-siber-guvenlik-stratejisi-ep-2020-2023.pdf

United Kingdom. (2024, May 14). *National Cyber Security Centre*. Https://Www.Gov.Uk/Government/Organisations/National-Cyber-Security-Centre.

United Nations. (2024, June 15). *Artificial Intelligence*. Https://Www.Unesco.Org/En/Artificial-Intelligence.

United Nations. (2024a, June 13). *UN Global Pulse*. Https://Www.Unglobalpulse.Org/Un-Global-Pulse/.

United Nations. (2024b, June 16). *High-Level Advisory Body on Artificial Intelligence*. Https://Www.Un.Org/Techenvoy/Ai-Advisory-Body.

Uwagbole, S. O., Buchanan, W. J., & Fan, L. (2017). Applied Machine Learning predictive analytics to SQL Injection Attack detection and prevention. *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, 1087–1090. https://doi.org/10.23919/INM.2017.7987433

VectraAI. (2024, May 7). *Vectra AI*. Https://Www.Vectra.Ai/.

Veeramachaneni, K., Arnaldo, I., Korrapati, V., Bassias, C., & Li, K. (2016). AI^2: Training a Big Data Machine to Defend. *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)*, 49–54. https://doi.org/10.1109/BigDataSecurity-HPSC-IDS.2016.79

Vidovic, K., Tomicic, I., Slovenec, K., Mikuc, M., & Brajdic, I. (2021). Ranking Network Devices for Alarm Prioritisation: Intrusion Detection Case Study. *2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1–5. https://doi.org/10.23919/SoftCOM52868.2021.9559086

von Solms, B. (2006). Information Security - The Fourth Wave. *Computers and Security*, *25*(3), 165–168. https://doi.org/10.1016/j.cose.2006.03.004

von Solms, S. H. (Basie). (2010). The 5 Waves of Information Security – From Kristian Beckman to the Present. In V. and W. C. Rannenberg Kai and Varadharajan (Ed.), *Security and Privacy – Silver Linings in the Cloud* (pp. 1–8). Springer Berlin Heidelberg.

Williams, L., McGraw, G., & Migues, S. (2018). Engineering Security Vulnerability Prevention, Detection, and Response. *IEEE Software*, *35*(5), 76–80. https://doi.org/10.1109/MS.2018.290110854

Wu, T., Liu, S., Zhang, J., & Xiang, Y. (2017, January 30). Twitter spam detection based on deep learning. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3014812.3014815

Xu, T., Goossen, G., Cevahir, H. K., Khodeir, S., Jin, Y., Li, F., Shan, S., Patel, S., Freeman, D., & Pearce, P. (2021). Deep Entity Classification: Abusive Account Detection for Online Social Networks. *30th USENIX Security Symposium (USENIX Security 21)*, 4097–4114. https://www.usenix.org/conference/usenixsecurity21/presentation/xu-teng

Yehezkel, A., Elyashiv, E., & Soffer, O. (2021). Network Anomaly Detection Using Transfer Learning Based on Auto-Encoders Loss Normalization. *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security*, 61–71. https://doi.org/10.1145/3474369.3486869

Yen, T. F., Oprea, A., Onarlioglu, K., Leetham, T., Robertson, W., Juels, A., & Kirda, E. (2013). Beehive: Large-scale log analysis for detecting suspicious activity in enterprise networks. *ACM International Conference Proceeding Series*, 199–208. https://doi.org/10.1145/2523649.2523670

Yen, T.-F., Heorhiadi, V., Oprea, A., Reiter, M. K., & Juels, A. (2014). An Epidemiological Study of Malware Encounters in a Large Enterprise. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 1117–1130. https://doi.org/10.1145/2660267.2660330

Zhang, L., Wang, S., & Liu, B. (2018). *Deep Learning for Sentiment Analysis : A Survey*. http://arxiv.org/abs/1801.07883

Zhang, X., Zhang, Y., Zhong, M., Ding, D., Cao, Y., Zhang, Y., Zhang, M., & Yang, M. (2020). Enhancing State-of-the-art Classifiers with API Semantics to Detect Evolved Android Malware. *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 757–770. https://doi.org/10.1145/3372297.3417291

Zhang, Y., Niu, J., He, G., Zhu, L., & Guo, D. (2021). Network Intrusion Detection Based on Active Semi-supervised Learning. *2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, 129–135. https://doi.org/10.1109/DSN-W52860.2021.00031

Ziolkowski, K., & NATO Cooperative Cyber Defence Centre of Excellence. (2012). *Peacetime regime for state activities in cyberspace : international law, international relations and diplomacy*.

Zscaler. (2024). *Zscaler Internet Access | AI-Powered Security Service Edge*.

ÖZGÜNLÜK BİLDİRİMİ

Uzmanlık tezi olarak sunduğum bu çalışmayı, bilimsel ahlak ve geleneklere aykırı düşecek bir yol ve yardıma başvurmaksızın yazdığımı, yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu, bunlardan her seferinde değinme yaparak yararlandığımı ve Bilgi Teknolojileri ve İletişim Kurumu Meslek Personeli Yönetmeliğine uygun olarak hazırladığımı belirtir, bunu onurumla doğrularım.

Bilgi Teknolojileri ve İletişim Kurumu tarafından belli bir zamana bağlı olmaksızın, tezimle ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara katlanacağımı bildiririm.

29.06.2024

Doğukan Ömer Gür

ÖZGEÇMİŞ

1993 yılında Adana'da dünyaya geldi. 2018 yılında İhsan Doğramacı Bilkent Üniversitesi Bilgisayar Mühendisliği bölümünden mezun oldu. 2019 yılında yedek subay olarak askerlik görevini tamamladıktan sonra 2020 yılında göreve başladığı Bilgi Teknolojileri ve İletişim Kurumunun Bilgi Teknolojileri Dairesi Başkanlığında Bilişim Uzman Yardımcısı olarak çalışma hayatını sürdürmektedir.